

FIGURE 101

GGGCCC GACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANCACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAGGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCCTGCGA

702 1237

FIGURE 102

ATTCTCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTGTTGCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCTATGACCCCAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTTCCTGTCCC

103/237

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGACAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTCGNGGNTTCTATGACCCTATGACCCAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTGTCCCCGAA

704/237

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTGTACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGAACC

105/237

FIGURE 105

TCATAGGGGGGCGCGATATTTTTCTTGCAGGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCTATGACCCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCCTG

106/ 237

FIGURE 106

TTCCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

107/237

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGACGAGGGGCTGTGGATGTCNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCCTGCGAA

108/237

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCGCCCGTGCCCCG
ACCGGTCCCCGCCTTTTTGTAAACTTAAAGCGGGCGCAGCATTAAACGCTTCCCCGCCCGGT
GACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGCTAAGGAACATGGCGGAAGGTGGAG
CAGGTCCTGAGCCTCGAGCCGCAGCACGAGCTCAAATTCGGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTTGCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAAACTTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAATAAAATTATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTCTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAAGCAGTTCAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCTTGTAAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACAAATATAATGTAACGATCTTTTAGAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTTATTAGTACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTGGAGCCACGAT
GCTGGGGAGTGCGGTGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCAGGCTG
CTTTCCGTGTCTTCAGTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTAATCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAA
GCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAATGTTATTTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTACAAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACCTCTGTTGGGTGAAGTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGTTTTATTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTCTGGGTGAGTAAATAACAACGTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAAATGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGGTGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCCCCAAAATTAAGAAAA

109/237

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFSTDVVTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMLOPFDYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIIGKIAL

Important features:**Transmembrane domain:**

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

1701 237

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTTGTCAATTTGTCACATTTGCTCTATGGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTAACTTTAAAATGAGC

111/237

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAAGTGTATTTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

7121237

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

773/237

FIGURE 113

GGTGGCCCATTCCCGGCCAGGCTGCTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTCAANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACG
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

7-14/237

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGTGTATTTCAGAGATGTTTAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

115/237

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCCAANTCCC
GGCCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTcantt
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTTCATNTCATGTTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

776/237

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAA
CTGTTATTCAGAGATGTTTAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTT
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

777/237

FIGURE 117

CGGAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAGATCAAAATGGAAAAATATGCTGCTTTGGTTGATATTTTTTACC
CCTGGGTGGACCCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTCACTGAAAGGACTTTCCATCTCACCAGCCCCGCATTTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCCCTATGAGACTGTCTTTGAGAATGGCACCCGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTTCAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTTCAGCACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCCTCAGCATGTTCTAACTGCTGCCACTGTGTTTCATGATGGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAGAGAGGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGGCCGGGGTCT
AGAGGATTGCCGAAGGGAGGCCTTCCTTTTCAGTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAATGGAATCAGCCCAACGATCAAGAAAA
TGCTGGTGAATGATCCACTTCTCAGGATTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTTCGGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTTCAGAAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTACCGG
GAACGATGCCAATTGTGCTTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTT
GAACTCTGTCAATAGCATTTCACATTTTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTTGTACTA
CTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTGTGTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGCATTATAAA
CAAACTAATACTGTTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAAATATGGACTTTTCATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTATAGGACAAAGAATTCTGTAATCTTTTCAAGAAAGAGT
CTTTTCTCCTTGACAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTTTAGATATGTCCTTTCCTAAAAATGAATAAAATTTATGAATATGA

778/237

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLLIDGSEMEWDFMWWHLRKVPRIVSERTFHLTSPAFAEADAKMMVNVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQV
YGTDSRFSILDKRFLTNFPFSTAVKLSTGCSGILISPQHVLTAACHVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGGRRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGMGDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCSVSDENLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

779/237

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGGAATGCTGCCGTCCGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCCGGGGAGGGGCCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCTTGCTTAAAAGGTCCTGATCACTT
ATATCTGGAAACCAAACCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTTATTGTCAAGATTTCGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCCTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGGCATCCAGAGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTTCCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAGAGGAGCCCTCGTAAGTTGTAAAAGCACAGACTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAAGATTGATTAGTTTCAA
AAAAA

120/237

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRRATPGTLLLLFLAFLLLSSRTARSEEDRDGLWDAGPWSECSRTC GGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNPCSLKCQ
AKGTTLVVELAPKVLDGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDLHLYLETKTLQGTKGENSLSSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFP
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKL PVEAKLPWFKQAQEELEGA AVSEEPS

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

7271237

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTTCGTGGGGCCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAGTTTCAGCTCCCAGCA
GGATTGTGGTAGTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAAGTGGCTAACATTCTTTT
TACCAGGGAAGTACCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACATTCCACTGTTGGTCAAACCACTCTTC
AATTGGTGTGTCATGGGCTTTTTTCAAAGTCCAGTAGAAGGTGCCAGACTTCCATTTATTT
GGCCTCTTACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAGTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATTAGGAACAAGGAGTAAAAGAGCTGTTTATAAAAGTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTTCAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAATTTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELQAACGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNH LGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTVNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:**Signal peptide:**

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

123/237

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTTTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

724/237

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCCTAACCCAACCCAACCTAGCCCAGTCCCAGCCGCCAGCGCCTGTCCCTGTCACGGAC
CCCAGCGTTACCAATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCCTGTAACAAGTAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTGGAGGAAGCTTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCACCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGGAGCAAAAGGACTCGGACAACCTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTTGGGGATGTTTCAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGATGTGACTTACAATTGGATTCAAGATAAATGTGTTCCCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTTAGAAATATTCCAGAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTCG
TATTTGACTTACATTCTGGAAAACCTGCACAGAGAATTCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAAACTTG
AAAAACAGTTTGTAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

725/237

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMM KREYR
GQRSVKALADYIRQQKSDPIQEIRD LAEITTLDRSKRNIIGYFEQKDS DN YRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDV LIPGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:**Signal peptide:**

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

726/237

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAGCGATCAGTGAAAGCATTGGCAGATTA

727/237

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGGCTCCCGGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTTATGCTGACTGGTGTGCTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

728/237

FIGURE 128

GCCCACGCGTCCGATGGCGTTCACGTTGCGGGCCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCTTGTACTCCCAGAGTACCTCAT
CCACGCTTTTCTTCTGTGTCATGTTTCTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCCAGGA
CTCTATGACCCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTTTATCTTCTAGCATTTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAATTAATATAAAAT
GATTACCTCTGGTGTTGACAGGTTTGAAGTGCACCTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGGAAGCTTTTGTATAGGAACCTTGTA
GGGCTCATTTTGGTTTTATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTTCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTGTAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAGTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTTCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

729/239

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:**Signal peptide:**

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

730/237

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCTTGTA TCTCCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCCCTCTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

131/237

FIGURE 131

CGGACGCGTGGGGGAAACCCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCGGGGACCGCTTCGGCTGAAGCATTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCTAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACCTCTGGGTGGATTTTAACTACAACCTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAATTCCTCCTCATAGAGCTTTTAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGS LWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRAQQLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSMMDSAQSFITSSWTFYQLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:**Signal peptide:**

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

733 1237

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTCCG
TGAAGTGAAGACAAGAACAATTATGTCCCTGATGCCAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

734/237

FIGURE 134

CACACTGGCCGGATCTTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCCGCTGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTTCGGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCAGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCCGCGGCTG
GGGATTCTTGTTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGGAAAATATGGAATGTCATCTACGAAGAAACTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGATTTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCCAGATTTTCAACTCTTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG
CATTTTGATGAGAATTCATTTTTTGCTGGGGATAAAAAAGAAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACCTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGCATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAAC
TCAGGAACTTGTTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAGGAGAATTATATTGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTTAAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

736(237

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAARQCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVDPGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPPFFERPDLQFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

737/237

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTTACCATT
CTTNGAGCGCCCAGATTTTCACTNNTTACTGGAAATAAAATTCAGGATGAGGNAAACAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTTCCTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

138/237

[illegible]

739/237

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGGFMVS

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

740/237

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCCTTCCAACCTTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAACTGGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

747/237

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTGCAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCCTGGTACTGGCCCTGTTTGCCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCCTCCAGACACCTTGAAAATAACCAATTACCCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACCTCTCAGGGCCTGGAT
CTCA**TAG**GTTTGCAGAAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAATACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAA

742 1237

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHFPNRRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLV LALFAFVGFM LILV
VVPLFVWKMGRLLOYS CCPVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

143/237

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

744/237

FIGURE 144

CCCACGCGTCCGCCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGAGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGGACAGAGGAGCTGAGAGAGGGGGGCGAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCATTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAGTGCAGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACCAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAAGTCTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCATTCTCCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCCACAGGGCACTGAGGGGC
AACAGGGACCCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGAGACTCCCCTTCGAGGATTGCACCCGCCCCGTCTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVAOSTNPFLSRLLNRDTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
VTWILIDRALNITSLQMHSRLRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

746/237

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTGCGCGCCAGCCTCCGCGCCGAGCCTC
GTTTCGTGTCCTCGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTTCTTCAGGTTGAGGGAGC
CGCAGAGGCGGAGGCTCGCGTATTCTGTCAGTCAGCACCCACGTGCCCCCGGACGCTCGGTGCTCAGGCCCTTC
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGGCCGTCCGGTTTGGCT
CACCTCTCCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
ATCCTGAGGTCAATTCATTATGAAGTGTACCGCGCGGAGTGGCTCAGAGTAACCCACAGTGCTGTTTCATGGCTAGA
GCAATTCAGCCATGGTGGTTCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCCAAACAACGAGGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGTTTCATGTACAATCGTGGTATGATGAAGTGAAAGACTTT
AGTACCCCATATGAACATGAATGCAACCCATATTGTCCATTACAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTGCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGCATAACATGAACATCTGGGGGCGAG
ATATGGCCCAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAGGGAAGTGGTGGGGCCATGCCCTTACAAA
CATGGGCGGCCCTGTCTGCTTGCCACCTAGTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAAGAGGG
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCAATAGCGCACAGCAATGTCCCAAATTTGTTTCT
TGTGAAGTAAGATTAAAGAGATCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCTGCTGGCTGGTTTGGAT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTTCATTATGGT
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTG
ACTTGTGAAACAACGTGGGAACAGCTCTGTCCATTTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACTGTCT
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATTGGAACTCGAGTTTATTCTGATCTGTCCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA
AAGACCTACATTGCTTCTTTTCAGAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATT
AGAGTGTTTGCTGTTGTGTGAAGTGAATACTTGAAGAGGACCATAAAGACTATTTCCAAATGCAATATTTCTGA
ATTTTGTATAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTCAGCCCAAAAAGGTGCCAAATGCATA
TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
GTTCTACGTTTCATATATTATATGGTGCTTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACCTGAAAGAAACCTTATCACATTTTCCCC
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC
TGTTAATTTAGGCATATAGAATATTAATTTCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC
CAAATGAATCTGTTAAATGTTTGATTCCCTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTCTATGTACTGTTA
AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
GAACAAAGATGAACATAATGTATTACATTACCATTTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA
AATCTGTAATGTTAGTTTTGGTAATTTTTTTTCTGCTGGTGGATTTACATATTAAATTTTTTCTGCTGGTGGG
TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRTVTTVLFMARAI PAMVVPNATLLEKLLLEKYMDEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQV VWATSNRIGCAINLC
HNMNIWGQIWPKAVYLV CNYS PKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTT CNR
YEC PAGCLDSKAKVIGSVHYEQSSICRAAIHYGIIDNDGGWVDITRQGRKH YFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFH KPA SHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV
FAVV

Important features:**Signal peptide:**

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

748/237

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGGCCGCGCTCCGACGGGGCCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCC
CGCCCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGGAAAACCTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTT
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACCTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTTGCTCCATTCCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAATGTGCTTCATTCCCCCT
GGTTAATTTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAAC
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCC
CAGTATATGCCGCATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAAAA

749/237

WO 00/53756

PCT/US00/04341

FIGURE 149

MSLLPRRAPPVSMRLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNANNEKRRVYEE

Signal sequence:

amino acids 1-34

750/237

FIGURE 150

GCCCCAGGGACTGCTATGGCTTCCTTTGTTGTTACCCCCGGTCTGCGTCAATGTTAAACTCCAATGTCCTCCTGTG
GTTAACTGCTCTTGCCATCAAGTTCACCCCTCATTGACAGCCAAGCACAGTATCCAGTTGTCAACACAAATTATGG
CAAAATCCGGGGCCTAAGAACACCGTTACCCAATGAGATCTTGGGTCCAGTGGAGCAGTACTTAGGGGTCCCCCTA
TGCCTCACCCCCACTGGAGAGAGGCGGTTTCAGCCCCCAGAACCCCCGTCTCCTGGACTGGCATCCGAAATAC
TACTCAGTTTGCTGCTGTGTGCCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCATCTGGTT
TACCGCCAAATTTGGATACTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCCTTTACTTAAACATCTACGT
GCCCACGGAAGATGGAGCCAACACAAAGAAAAACGCAGATGATATAACGAGTAATGACCGTGGTGAAGACGAAGA
TATTATGATCAGAACAGTAAGAAGCCCGTCATGGTCTATATCCATGGGGGATCTTACATGGAGGGCACC GGCAA
CATGATTGACGGCAGCATTTTGGCAAGCTACGGAAACGTCATCGTGATCACCATTAACACCCTGCGGAATACT
AGGGTTTTTAAAGTACCGGTGACCAGGCAGCAAAAGGCAACTATGGGCTCCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAAATGTGGGAGCCCTTTGGCGGGGACCCCAAGAGAGTGACCATCTTTGGCTCGGGGGCTGGGGCCTC
CTGTGTGACGCTGTTGACCCCTGTCCCACTACTCAGAAGGTCTCTTCCAGAAGGCCATCATTAGAGCGGCACCCG
CCTGTCCAGCTGGGCAGTGAACCTACCAGCCGGCCAAGTACACTCGGATATTGGCAGACAAGGTGCGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCCC
GGCCACCTACCACATAGCCTTCGGGCGCGGTGATCGACGGCGACGTCATCCAGACGACCCCGAGATCCTGATGGA
GCAAGGCGAGTTCTCAACTACGACATCATGCTGGCGGTCAACCAAGGGGAAGGCCTGAAGTTCGTGGACGGCAT
CGTGATAACGAGGACGGTGTGACGCCCCAACGACTTTGACTTCTCCGTGTCCAACCTTCGTGGACAACCTTTACGG
CTACCCTGAAGGGAAGACACTTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGGCCGATAAGGAAACCC
GGAGACGCGCGGGAACCCCTGGTGGCTCTCTTTACTGACCACAGTGGGTGGCCCCCGCGGTGGCCGCGGACCT
GCACGCGCAGTACGGCTCCCCACCTACTTCTATGCTTCTATCATCACTGCCAAAGCGAAATGAAGCCAGCTG
GGCAGATTTCGGCCCATGGTGATGAGGTCCCCTATGTCTTCGGCATCCCCATGATCGGTCCCACCGAGCTCTTCAG
TTGTAACTTTTCCAAGAACGACGTCATGCTCAGCGCGGTGGTCATGACCTACTGGACGAACTTCGCCAAACTGG
TGATCCAAATCAACAGTTCTCCTCAGGATACCAAGTTCATTACACAAACCCCAACCGCTTTGAAGAAGTGGCCTG
GTCCAAGTATAATCCCAAAGACCAGCTCTATCTGCATATTGGCTTGAAACCCAGAGTGAGAGATCACTACCGGGC
AACGAAAGTGGCTTTCTGGTTGGAACCTCGTTCTCTATTGCACAACTTGAACGAGATATTCCAGTATGTTTCAAC
AACCACAAAGGTTCTCCACCAGACATGACATCATTTCCTATGGCACCCGCGGATCTCCCGCCAAGATATGGCC
AACCACAAACGCCCAGCAATCACTCCTGCCAACAACTCCCAAACACTTAAGGACCCTCACAAACAGGGCCTGA
GGACACAACCTGTCTCATTGAAACCAACGAGATTATTCACCGAATTAAGTGTCAACATTGCGGTGCGGGCGTCT
GCTCCTCTCTCTCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAAAGGACAAGAGGCGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGCTCACATCCAGAACGAAGAGATCATGTCTCTGCAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTGCCTGCAGGCACACGACACACTGAGGCTCACCTGCCCCGAG
CTACACCCTCACGCTGCGCCGGTGCAGATGACATCCCACTTATGACGCCAAACACCATCACCATGATTCCAAA
CACACTGACGGGGATGCAGCCTTTGCACACTTTTAAACACCTTCAGTGGAGGACAAAACAGTACAAATTTACCCCA
CGGACATTTCCACCCTAGAGTATAGCTTTTGCCCTATTTCCCTTCTATCCCTCTGCCCTACCCGCTCAGCAACAT
AGAAGAGGGGAAGGAAAGAGAGAAGGAAAGAGAGAGAGAAAGAAAGTCTCCAGACCAGGAATGTTTTGTCCCACT
GACTTAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGGCAGACCCTTATCGTTGGTGTTCAGTATTAC
AAGATCAACTTCTGACCCTGTGAAATGTGAGAAGTACACATTCTGTAAATAAAGTCTTTAAGATCTCTACCA
CTCCAATCAATGTTTAGTGTGATAGGACATCACCATTTCAGGCCCCGGGTGTTTCCAACGTCATGGAAGCAGCT
GACACTTCTGAACTCAGCCAAGGACACTTGATATTTTTTAATTACAATGGAAGTTTAAACATTTCTTTCTGTGC
CACACAATGGATGGCTCTCCTTAAGTGAAGAAAGAGTCAATGAGATTTTGGCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTTATTATTAAGAATGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAGAG
GTGTTTTGCCAGCCTGAACATATTTAAGAGACTTTGT

757/237

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVVPYASP
PTGERRFQPPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTSLHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPSWADSAHGDEVYPYV
FGIPMIGPTELFSCNFSKNDVMLS AVVMTYWTNFAKTGDPNQVPVQDTKFIHTKPNRFEEVA
WSKYNP KDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPDMS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSDPHKTGPEDTTVLIETKRDYSTEISVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

152/237

FIGURE 152

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTTGTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTTCGAGTACTTGAAACGGGAGCACTCGCTGTCTGAAGCC
CTACCAGGGTGTGGGCACAGGCAGTTCCTCACTGTGGAATCTGATGGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCTTACCCCAGATATGCAAAGTAAACAGGGTGCCTTGTGGAACCGG
GTGCCATGTTTCTTGAGAGACTGGGAGTTGCAGGTGCACTTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCCTACATCTCAGCCATGGTGAACAACGGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCTACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCCTGGTGATTCTGCTACGTCAAGAGGCATTTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAAGTGACAGTGGAGAGAACCCCAAGAGGAAAAGCTCCATCGA
GATGTGTTCTTGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCCTGGCCCTCTTCCTCATCGTCTTTTCTCCCTGGTGTGTTTCTGTATTTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACTTTTGTGACTGTCACCCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTTCTTGCTCTCTAGCAGCTGGTTGGGACTATATTCTG
TCACTGGAGTTTGAATGCAGGGACCCCGCATTCCCATGGTTGTGCATGGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCTTTCCCTGCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAAATCACAGAAC
AGAATTTCATAGCCCAGGCTGCCGTGTTGTTTGAAGTCAAGAGGCCCTTCTACTTCAGTTTTG
AATCCACAAAGAATTAAAACTGGTAACACCACAGGCTTTCTGACCATCCATTCGTTGGGTT
TTGCATTTGACCCAACCCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAACT
TCTTCCCTGCCTTACCTTCCTTTCCTCCATTTCATTGTCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAATAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGATCAAACCAAAAGCAACATTTGTGATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGT
TACGATTTTTGGAATCCCACTTTGAGTGCTGAAAGTGAAGGAAGCTTTCTTCTTACACCTT
GGGCTTGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTT
CACTGTCAATTGAGCAGCACAGCCTGAGTGCTGGCCTCTGTCAACCCTTATTCCACTGCCTTA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAG
AGTCTCCTTGGAGGGCCTGGAACCTCTGAGTCCTCCTATGAACCTCTGTAGCCTAAATGAAAT
TCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCGGCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATG

753/237

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLLGSGQGPGQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKEIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFELPSVDNMKLPMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

754/237

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAAAATCTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGAGCCCCCTGTGA
CAGACCCCGTTTATGAAGCTCTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTGATGCC
CGCATCATTTTAAGCTGGTCTCAGTGCAATGTTTCATTGCGCCACGGAGACAGGTACCCACTGTATGTCATTCCCA
AAACAAAGCGACCAGAAATTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTCA
TTAGTCACATGTCAAAGGATCCGGAGCCTCTTTCGAAAGCCCTTGAACCTTGCCTTTACCCAAATCACC
CATTGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAACTCCTGCCAATGATTGGTCTGCAGACCAGCTCTATTTAGAGACCCTGGGAAAA
GCCGACCCCTACAAAGTGGGCTGGCCTTGCTTTATGCTTCTCCAGATTTTGACTGGAAGAAGATTTATTCA
GGCACCAGCCAAGTGCGCTGTTCTGCTCTGGAAGCTGCTATTGCCCGGTAAAGAAACCATCTGGAAAGGAGC
AGCGTCGTACGTACCTCCTACGTTTGAAGAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCTAGACTCCATGCTCTGCCACTTCTGCCACAATGTCAGCT
TTCCCTGTACAGAAATGGCTGTGTTGACATGGAGCATTCAAGGTAATTAAGACCATCAGATCGAGGATGAAA
GGGAAAGACGGGAGAAGAAATTGTACTTCGGGTATTCTCTCCTGGGTGCCACCCCTCCTGAACCAACCATCG
GCCGATGCGAGCTGCCACCGAGGGCAGGAAAGAGCTCTTTGCCCTCTACTCTGCTCATGATGTCAGTCTGT
CACCAGTTCTCAGTGCCCTTGGGCTTTTGAAGGCCAGGTTCCCAAGGTTTGACGCCAGGTTGATCTTTGAGCTTT
GGCAAGACAGAGAAAAGCCAGTGAACATTCCGTCCGATTCTTACAATGGCGTCGATGTCACATTCACACCT
CTTTCTGCCAAGACCACACAAGCGTTCTCCAAGCCATGTGCCGCTTGAAGAACTTGGTCCGCTTTGTGAAAA
GGGACATGTTTGTAGCCCTGGGTGGCAGTGTTACAAATTATTATGATGCATGTCACAGGGAAGGATTTCTAAAGG
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAGGTCCACTTCTAGTTTTGTCTGTAC
TAAGGGTAGAAGATTATTGCTTTTTAAAGGCTAAATATTGTTTGTGGGAACACAGATGGTTGGGTTGAACAGT
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGTACTT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTGCAC
TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACAGCACACCTCAACCAAAATTTTTTAATCTTAGACATT
TTTACCTTGCTCTGTTAAGAAATTTCTGAAGTGATTTATCTAAAATAAAGGTTGGCAAACTTTTTCTGTAAAGG
GCCAGATTGTAAATATTTAGACTGTGTGGACCAAAAGGCCACATACAGTCTCTGTCTAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCAGGCCAGACAAAACA
GATGGTGACCAGACTTGGCCCTGGGCTGTAGTTTGTGACCCCTCATCTAAAAATAGGCTATACTACAATTGC
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATTCAATGGTTCTCCAGTAACCTTCTGTAGAAACA
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTTGTTACAAGATAGTTTTTAAGTATGTTCTAAATATTTGT
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTGTATGCCATTTAGTATTTTATAGTTTAGGAAAAATTT
TTCTAAGACCAGTTTTAGATGACTCTTATTCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTCAGGCACTTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA
ACTGGATTCAATTTTAAACCAATTTTCATCAGTTTCAAATGGTAAATCTGATTGATTTTTAAATGCGTTTTTGA
AGAATTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTTCTGAATAATGGTGCTAATTCAGAGAAATGGAAAGTGAAGTGAGATTCTCTGTTGTCATCGGCATTCC
AACTTTTTCTCTTTGTTTTGTCCAGTGTTGCATTTGAATATGTCTGTTTCTATAATAAATTTTTTAAGAATAA

155/237

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLLALAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDVY
EALLYCNIPSPAERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSHSVRILYNGVDVTFHTSF
CQDHHKRSPPKPMCPLENLVRVFKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

156/237

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCATCCCCCTTTGAAGAACAGTACTGTGGA
GCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC
GCCGCTGTTTACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGACGCGCCCTGAAGCACAAAGCAGAT
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAACAACGGAGGAGCTCTGACTTCCCAACTGTCCATTCTAT
GGGCGAAGGAACTGCTCCTGACTTCAGTGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT
TCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
ATGCACCCAGATACGCTATTTCAGTTCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCT
GAATCCGCGCAGCGCGAGCTTGGTCACGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG
TCAATTAATCTAGACATTCTGATGGAGGATAAAGTGAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTTCTGTAAAGTGAATTAGAAATAAAAATTAGTGAATATGCAGCCACTGAGATGCGGTT
CCCTCTACCCACGCTGGGATCCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCCGAACACTCA
CTTCTCCCTCATCGTGAAGTGGACCGGAGCGGAGTACCGCGAATTGGTGTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTCACCACCTGGTCTTACGGCTCCGACGGGGGCGACCCGGTGCACAGGCACCGCGCG
CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCGCTTTGCTCAGCCCGAGTACCGCGCGAGCGTTCC
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT
GAGTATTCCCTCCGGTATGTGGACGACAAGCGCGGCCAAGTTTCAAAGTATGTAATTGAGGACAAATATC
AACAATAGGGGAGTTGGACCACGAGGAGTCAGGATTCTACCAGATGGAAGTGAAGCAATGGATAATGCAGGATA
TTCTGCGCGAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCGAGAAGTGGTCTCACCTCTCT
CGCCAGCTCGGTTCCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAATGACCAAGATTCTGA
GGAACCGGAGTGTGATCTGTTTCAATCCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTAATA
TAGTTTAGTCACAGACATAGTCTTGGATAGGGAACAGGTTTCTAGCTACAACATCACAGTGACCGCCACTGACCG
GGGAACCCCGCCCTATCCACGGAACTCATATCTCGTGAACGTGGCAGACACCAACGACAACCCGCGCGGTTCT
CCCTCAGGCCTCTATTCCGCTTATATCCAGAGAACAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA
CGAGCCGACTGTGAAGAGAACGCCAGATCACTTATCCCTGGCTGAGAACACCATCCAAGGGGGCAAGCCTATC
GTCCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCGAGGA
CTTGCAAGTGAAAGTGATGGCGCGGGACACGGGCACCCGCCCTCAGCAGCAACGTGTCGTTGAGCCTGTTGCT
GCTGGACCAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCCTCCCCACGGACGGTTCCACTGGCGTGGAGCT
GGCTCCCCGCTCCGCGAGAGCCCGGCTACCTGGTGACCAAGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
CTGGCTGTCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGCCGCTCAGGACACCGGCCAGCC
CCCTCTCTCCGCCACTGTACGCTCACCCTGGCCGTGGCCGACAGCATCCCCAAGTCTGGCGGACCTCGGCAG
CCTCGAGTCTCCAGTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCCGGCTCTCTG
CGTCTTCTGCGCTTCTGTCATCTTGTGCTGGCGCTCAGGCTGCGGCGCTGGCACAAGTCACGCCTGCTGCAGGC
TTCAGGAGGCGGCTTGACAGGAGCGCCGCGCTCGCACTTTGTGGCGGTGGACGGGTGCAGGCTTTTCTGACAG
CTATTTCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACCTGATCTTCCCCAGCCCCAACTATGCAGA
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAAGCGAGCCCTTTTGTGTGTCAGGTGATTGCGGATTTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAATCTTCTTTCTTTTTTTTTTAATTGCTCTGTCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTACTGCGGCCTCAAACCTCCTAGGCTCAAGCAATTATCCCACCTTTGCT
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCTATCTATCTATCTATCTATCTATCTATCT
CTATCTATCTATCTATCTATTACTTTTCTGTACAGACGGGAGTCTCACGCCTGTAATCCCAGTACTTTGGGAGGC
CGAGGCGGGTGATCACCTGAGGTTGGGAGTTTGGAGCCAGCCTGAACCAATGGAGAAACCCGCTCTATACTAA
AAAAATACAAATAGCCGGGCGTGGTGGTGCATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT
TGCTTTAACCCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCCTCCAGCCTGGGCAACAAGAGT
AACTCTATCTCA

157/237

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVP EELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFFLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYISLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLIALSSFDYEQFRDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLESPA
NSETSDLTLYLVVAVAAVSCVFLAFVILLIALRLRRWHKSRLQASGGGLTGAPASHFVGVD
GVQAF LQTY SHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVS LYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

758/237

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCTT
ACAGCTGCCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

759/237

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

760/237

FIGURE 160

GGCGCCGGTGACACGGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCTGCGCGCCCCGGCCCCG
CGCGCCGCCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCGGGCCCCGCGCCCCG
GCCCCGCGCCAGGTGAGCGCTCCGCCCCGCGCGAGGCCCCGCCCCGGCCCCGCCCCGCCCCG
CCCCGGCCGGCGGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACCTGATCCCATAAAAC
ATTCATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCGCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCCAGCCAGAGCCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGGCCGGGCCGGGGCCGGGCCGTAGCGGCGGCGCCTGGA
TGCGGACCCGGCCGCGGGGAGACGGGCGCCCGCCCCGAAACGACTTTCAGTCCCCGACGCGC
CCCCGCCAACCCCTACG**ATGA**AGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAGGTGACGACAAGCTGCCCCCAGCAGGGCCTGCAGGCTGTGCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGACGCGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCGCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCGAATTGATGC
GGCTGCCTTCACTGGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCTGCCACATTCCACGGCCTGGGCCGCTACACACGCTGCACCTGGACCGC
TGCGGCCCTGCAGGAGCTGGGCCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCA
CACACCTCTTCTGACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAGAACCGCGTGCCCCATGTGCACCCGCATGCCTT
CCGTGACCTTGGCCGCCTCATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCCACTCTGGGCCCTGGCTGCAGAAGTTCCGCGGCTCCTCCTCCGA
GGTGCCCTGCAGCCTCCCGCAACGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGT
ACTGGAGCCTGGAAGACCAGCTTCGGCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCCGGTG
ACAGCCCCGCCGGGCAACGGCTCTGGCCCACGGCACATCAATGACTCACCTTTGGGACTCTG
CCTGGCTCTGCTGAGCCCCCGCTCACTGCAGTGCGGCCCCGAGGGCTCCGAGCCACCAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGGCCAGGCTGTTACGCAAGAACCGCACCCGCAGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCT**TGA**CCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCCTCCCTGATGGACGCCTGCCGCCCGCCACCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCAGAACGCCGCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAA
AAAA

7611237

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLGELGPGFLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCDRCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPWIWTGRATDEEPLGLPKCCQPDAAKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGSALPSLTCSLTPLGLALVLWTVLGPC

Important features:**Signal peptide:**

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

162/237

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCCTCTGGAGAGGACTAC
TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTAC
AGTCCACAGAACCGTCCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCGCAGCGGGA
AGCTCATTTGCAGACAAAGGCAAGTCCTTTTTTCTCTCTCTTTGGGCTTATCTCTGGCG
GGCGCGGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTTGTCAC
CAATTTAGCAAAGGACCTGGGTCTGGAGCAGAGGGAATTCTCCAGGCGGGGGGTTAGGGTTG
TTTCCAGAGGGAACAACTACATTTGCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTCACACAGAGCCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAAACAAATGTGGTGAAAGTATCAGAGAGCAGTCTCTCTGGG
ACTACGTTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTCGGGTCCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCCCTGAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCTCTGGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGGAGAGATTTCTTACTTTTCCAAGCTTCAGAAGAGATTGGCAAACCTTTAAGAT
CAATCCCTTGACAGGAGAAATTGAACTAAAAAACAACCTCGATTTGAAAAACTTCAGTCCT
ATGAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAAAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGTCTGCATTTACCAGCCCAAT
ACCTGAGAACGCGCCTGAACTGTGGTTGCACTTTTTCAGTGTTTCAGATCTTGATTTCAGGAG
AAAATGGGAAAATTAGTTGCTCCATTTCAGGAGGATCTACCCTTCTCTGAAATCCGCGGAA
AACTTTTACACCCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTCACCCAAACCTCCTACACCCTGTTCTGTC
CGCGAGAACAACAGCCCCGCCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCCAGGTACCTACTCGCTGCTGCCGCCCCAGGACCCGCACCTGCCCCCTCACAT
CCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCTGCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCAGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGCGTGGTGGTGTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCGGGCGGCGGAGCCGGGCTAC
CTGGTGACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTCTGTACCA
GCTGCTCAAGGCCACGGAGCTCGGTCTGTTCTGGCGTGTGGGCGCACAAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGCGACGCGGCCAAGCACAGGCTGGTGGTGTGTTCAAGGAC
AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGGTGGACGGCTTCTC
CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGACCCAGGCCCAGGCCGACTTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGTCTCTGTTT
GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCTCGGTGGGTGCTGCTTGGTGCCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACAGGACCCCTATCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCCTGAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
TAACTTTGGGTTCAATATTCAGTGAACCATAGTTGACTTTTACATTCCATAGGTATTTTATTT
TGTGGCATTTCATGCCAATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTTCTCATGTTCTTCTCCCTTGTGTTTAAAGTGAACATTTACCTTTATT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLGLSLAGAAEPRYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTPCVLRQVLLESFFEFFQAELOV
IDINDHSPVFQDKQMLVKVSESSPPGTTFFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLVDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTPLMITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELGFLGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFVSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMSGTRT
LSQSYQYEVCLAGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFENIQ

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

769/237

FIGURE 164

ACCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACC
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCCTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCCTCACTTTAACTCTCTGCCCCGGGCATTTCCAGCTCTT
CACTTTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCGTAGC
TGTTCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACCTCAAGCCGACCAAATAGGCCCTCTTCCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGATTTTCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTCGGTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCATAAGCAAATGC
AAAAATATTCAATAG

165/237

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRPRPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETCLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEHEVE

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296**Thioredoxin domain**

amino acids 211-227

766/237

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGGCCAGCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCCGACATGCTTGAGTGCGACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAAATAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTACAAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCACTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGCTCCCGTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCACCCGAATCTTGTAGAAATATTCAAACCTAATA
AAATCATGAATATTTTAA

767/237

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFS YKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLD ETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLNGVSE RDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQG GELVITSVKRWQKGQREFKRISRSIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

768/237

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTGTCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACCTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTTGCCTTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGCAACCCTCTGGGCGTCCTTGTGGCCAATGTGCT
GTCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATAACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACCCTCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCAGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTCTGA
GTGTTCCCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATAACGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGGCCGGGAGCCCCCACCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTTGTAAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

7691237

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVGTON
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNPLGVLVANVLSPLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFSALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCFLSLACVPFALVSQQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGLIMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPGRLAGRVQASRFIDPAGSHSSSFSSPWVIT

Important features:**Signal peptide:**

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

770/237

FIGURE 170

GTCCACATCCTGCTCAACTGGGTGAGTCCCTCTTAGACAGGCTCTGTGCATCATTGTGGAAGTGGACCAAC
TAGTTCCCCAGTAGGGGGTCTCCCTGGCAATTCTTGATCGCGTTTGGACATCTCAGATCGCTCCAATGAAGA
TGGCCTTGCTTGGGGTCTGCTGTTTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGCGAGCTCTGGGGG
AAGGAGCAGGGGCTGATCAAGCCATCCAGGAACACTGGAGGACTTGTCCAGCTTGAAGAACTCTAGTGGTT
TCTGAATCTAGCCCACTTGGCGGTAAAGCATGATGCAACTTGTCAACTTCTGTGGGGCTTTTGGGGCCAGGTGG
CTACTTATTTCTTTTAGGGGATTGTGAGGAGGTGACCACTCTCACGGTGAAATACCAAGTGTGAGAGGAAGTGCC
ATCTGGTACAGTATCGGGAAGCTGTCCAGGAAGTGGGCCGGGAGGAGAGGCGAGGCAAGCTGGGGCCGCTT
CCAGGTGTGAGCTGCCTCAGGCGCTCCCCACTCAGGTGACTTGTAGGAAGGCTGTCTCAGCACAGGAGGGC
CTGTGATCGAGAGCAGTGTGCCGACAGTGGGATCCCTGCTGGTTTCTTTGATGTGCTTGGCACAGGGGATTT
GGCTCTGATCCATGTGGAGATCCAAGTGTGACATCAATGACCACCAGCCAGGTTTCCCAAAGGCGAGCAGG
GCTGGAAATCTCTGAGAGCGCTCTCTGCGAACC CGGATCCCCCTGGACAGAGCTCTTGACCAGACACAGGCC
TAACACCCCTGCACACCTACACTCTGTCTCCCAGTGAGCACTTGTCCCTTGGATGTCTTGTGGGCCCTGATGAGAC
CAACATGCAAACTCATAGTGGTGAAGGAGCTGGACAGGGAATCCATTCTATTTTTGATCTGGTGTAACTGCT
CTATGACAATGGGAACCCCCCAAGTCAGGTACCAGCTTGGTCAAGGTCAACGTCTTGGACTCCAATGACAATAG
CCCTGCGTTTGTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT
GACCGCCACAGACCTTACCAAGGCCCAATGGGAGGTGGAGTTCTTCTCTAGTAAGCACATGCTCCAGAGGT
GCTGGACACCTTCAGTATTGATGCCAAGCAGGCCAGGTCATTCTCGCTGCACCTAGACTATGAAAAGAACCC
TGCTTACGAGGTGGATGTTGAGGCAAGGACCTGGGTCCCAATCTATCCAGGCCATTGCAAAGTTCTCATCAA
GGTTCTGGATGTCAATGACAACATCCAAGCATCCACGTACATGGGCCTCCAGGCCACTGGTGTGAGAAGC
TCTTCCCAAGGACAGTTTTATGTCTTTGCTATGGCAGATGACTTGGATTGAGGACACAAATGGTTTGGTCCACTG
CTGGCTGGACAAGAGTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAAACAAATGC
CACACTGGACAGAGAGCAGTGGCCCAAATATACCCTACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
AGCCAAAGAACAGCTCAGCATCTCAGATCAGTGACATCAACGACAATGCACCTGTGTTTGAGAAAAGCAGGTATGA
AGTCTCCACGCGGGAAAACAACTTACCCTCTCTTCACTTACCATCAAGGCTCATGATGCAAGCTTGGGCT
TAATGAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACTTAGTAGCTTAGTACTCCAACACAGGAGA
GGTCACTGCTCAGAGGTCACTGAACTATGAAGAGATGGCGGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTGAGCCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
CCAGCTGTGCTCAGCGATGGAAGAACCGAGCTCTCCGTGCTTGTGAATGCCTCAGCACAGGCCACCTGTGTGGC
CATCGAGACTCCCAATGGTTTGGGCCAGCGGGCACTGACACACTCCACTGGCCACTCAGAGTCCCGGCCATT
CCTTTTGACAACCAATTGTGGCAAGAGATGACAGCTCGGGGGCAATGGAGAGCCCCCTACAGACCTCCGCAATGG
AAATGAAGCCACCTCTTCATCTCAACCCTCATACGGGGCAGCTGTTCTGTCATGTCAACATGCCAGCAGCCT
CATTTGGGAGTGAGTGGGAGCTGGAGATAGTAGAGGACAGGGAAGCCCCCTTACAGACCTCGAGCCCTGTT
GAGGCTCATGTTTGTACCAGTGTGGACCACTGAGGACTCAGCCCGCAAGCTGGGGCTTGGACCTTGGAGTGTGAT
GCTGACGTTGATCTGCCTGGCTGTACTGTTGGGCATCTTCGGGTTGATCTTGGCTTGTGTCATGCTCATCTGCGG
GACAGAAAAGAAGGACAACAGGGCCTACAACCTGTGGGAGGCGGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
CCAGAAAACATTCAGAAAGGCAGACATCCACCTCGTGCTGTCTCAGGGGTGAGGAGGTGAGCCTTGTGAAGT
CGGCGAGTCCCACAAGATGTGGAACAAGGAGCGATGATGAAGCAGGCTGGGACCCCTGCTGACGGCCCCCTT
CCACTCACCCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACAGGAGGACCGGCGGAGAGCCGAGAGGT
GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCAGGCAGAGGAATGCCTCCCGGGAGAACCTGAAGCTTCC
CGAGCCCCAGGCTGCCACAGGCCAGGCCAGCTTCCAGGCTCTGAAGTTGCAAGGACGCCCAAGGAGGCTGGC
TGGAGACAGGGCAGTGAGGAAGCCCCACAGAGGCCACGAGCTCTCTGCAACCTGAGACGGCAGCGACATCT
CAATGGCAAAGTGTCCCTTGAGAAAAGATCAGGGCCCCGTGATGATCCTGCGGAGCCTGGTCCGGCTGTCTGTGGC
TGCTTTCGCGGAGCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCCTGTTGAGCAAATCTCCAGCTGTG
GTCCTTGTCTGATCAGGGCCAAATTCAGGCCCAAACCAAACCCAGGAGAAATAAGTACTTGGCCAAGCCAGGAGG
CAGCAGGAGTGAATCCAGACACAGATGGCCCCAAGTGCAAGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGTGTCAAGTCTGCTGGACCC
CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCCGTGACCCGGCTGGATGGCGAGACTCTCTTTGCCCTCAC
CACAACCTACCGTGACATGTGATCTCCCGGATGCTGCAGCCACGGAGGAGCGGAGGACCTTCAGACGTTCCG
CAAGGCAGAGGCCACAGAGCTGAGCCCCAACAGGCACGAGGCTGGCCAGCACCTTGTCTCGGAGATGAGTCACT
GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGGCTCTCGGT
CTGCGGGAGGACCCTCAGTTTAGACTTGGCCACCAGTGCAGCCTCAGGCATGAAAGTGAAGGGGACCCAGGTGG
AAAGCAGGGGACTGAGGCGAAGAGCAGAGGACGAGCAGCAGCAGCAGGTCCTTGAACATACCTCAGACGCCCT
CTGGATCCCAAGAACAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAATCTTGAATCTACTAGCTAG
CGGGCGCTTGAACCTTAGGGTGACTGATGCTACCCCAAGAGGAGGCAAGGCCAGGACTAACAGCTGAC
TGACCAAGCAGCCCCCTTGAAGCAGCTCTGAGTCTTTTGGAGGACAGGGACGGTTTGTGGCTGAGATAAGTGT
TCCTGGCAAAACATATGTGGAGACAAAGGGTCAGTCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
AAAGGCTGGCCTTCTTGGGTAGCAGGAGTCAAGGGCTGTACCTGGGGGTGCCAGGAATGCTCTTGACCTAT
CAATAAAGGAAAAAGCAGTAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGGAAAAAAGG

777 1237

FIGURE 171

```

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLTIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT
IVARDADSGANGEPLYIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLOAPFHLTPTLYRTLNRNQGNQGAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGS
RSAIPD TDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAASGMKVQGDPPGGKTGTGKSRGSS
SSSRCL

```

Important features:**Signal peptide:**

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

172/237

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATTGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTTCTGGATTTGGTGTGTCAACTGCCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGG
AAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTGGTCCCAACACATT
TCCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCCTCCAATGTCATTGTCCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGCCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTTAAGGTTACATGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

773/237

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTASGSENLTLIQ
QEVDALIELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:**Signal peptide:**

amino acids 1-23

Potential transmembrane domains:amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444**N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

7741237

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

775/237

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAATTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATATTTTTTGGATTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGCACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCATTC

796/237

FIGURE 176

CTCGCGCAGGGATCGTCCC**ATG**CGCGGGGCTCGGAGCCGCGACCCTTGGGGGGCTCCGGGATTGCTACCTTTT
TGGCTCCCTGCTCGTGAAGTCTTCTCACGGGCTGTCGCCTTCAATCTGGACGTGATGGGTGCCTTGGCGAA
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG
CCCGTTTGACCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA
GGAGAACCAGTGGTTGGGAGTCAGTGTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATA
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTGTGCTCAGCCAGGA
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT
TGGGTTTGCCAGCAGGGGCACAGCTGCCGCCCTTCTCCCTGATAGCCACTACCTCCTTTGGGGCCCCAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG
TCCCTACGAGCGGGGGGAGAGAAGGAGCAGGACCCCCGCTCATCCGGTCCCTGCCAACAGCTACTTTGGCTT
CTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAGAGCTGAGCTTGTGGCTGGAGCCCCCGCGCCAAACCA
CAAGGGTCTGCTGCTCATCTGCGCAAGGACAGCGCTCGGCTGGTGCCGAGGTGCTGCTGAGGTGCTGCTGAGGTGCG
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGCTGGCCAGACCTGATAGTGGG
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACAGGGGGGTCACTG
GGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGA
GCTCAACCAAGATGGCTTCCAGATATTGCAGTGGGTCCGCCCTTGTGATGGTGATGGGAAAGTCTTCATCTACCA
TGGGAGCAGCCTGGGGGTGTGCGCAAACTTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCTGACCTGCTGGTGGGCTCCCTGGCTGACAC
CGCAGTGCTCTTCAGGGCCAGACCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT
GGAGGGTCTGCTGCTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAAGTACATTCGAGTCCC
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT
TCCCCGTGTGACGTTCTGTGAGCCGTAACCTGGAAGAACCACAGCACCAGGCCTCGGGCACCCTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGCCATTGT
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCCTCCAGTGGC
CCCCATCCTCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCCAACC
TCTGCCCATGGATGTGGATGGAACAACAGCCCTGTTTGCAGTGTGGGAGCCAGTCAATTGGCTTGGAGCTGAT
GGTCAACCACTGCCATCGGACCCAGCCAGCCAGCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGT
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCTGCGGAGAAGCCACTCTGCCTGTCCAA
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTCACCTTCTACCTCAT
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAATGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
GGAGCTGATCCCATCTCTGCACGAGCCGTGTCTTCACTGAGCTGCCACTGTCCATTGACGGAATGGCCATTCC
CCAGCAACTCTTCTCTGCTGTGGTGTGGTGGGGGCGAGAGAGCCATGCAGTCTGAGCGGGATGTGGGCAGCAAGGT
CAAGTATGAGGTACGGTTTCCAACCAAGGCCAGTCCGCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCTGG
GCAGAAAGGGCTTTGCTCTCCAGGGCCCAACATCCTCCACTGGATGTGGACAGTAGGGATAGGAGCGCGCGGGA
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGAAGCGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGCTG
TGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGACGGCCAACCTGTGTGGTGTTCAGCTGCCCACTCTA
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA
TGCTCCACAGTGATCCAGTGATGGTATACTTGGACCCATGGCTGTGGTGGCAGAAGGAGTGCCCTGGTGGGT
CATCCTCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCACTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAACTGGGGCAGCCCCCGGGGAGGGCCGGATGCACA
CCCCATCCTGGCTGTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCT**TAG**GTCC
CATGTCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTGCTCATCAAGATTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAAC
TCCTCCACCCAACTTCCCCTTAGAGTGCTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG
TGAGAAGGGCAGGGGTGTCTGATGCAAAGGTGGGAGAAGGGATCCTAATCCCTTCTCTCCATTACCCCTGT
GTAACAGGACCCCAAGGACCTGCCTCCCCGGAAGTGCCTTAACCTAGAGGTCGGGGAGGAGGTTGTGTCACTGA
CTCAGGCTGCTCCTTCTCTAGTTTCCCCTCTCATCTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGT
TTCGTCTATTTATTAATAAATATTTGAGAACAATAAAAAAAAAAAAAAAAAAAAA

777/237

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDWPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGEWKFCFEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGPDLIVGAPYFFERQEEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLNDGFPDIAVGAPFDGDGKVFYHGSLSGVVAKPSQVLE
GEAVGIKSFYGSLSGSLDMDGNQYPDLLVGS LADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFYSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQ PSTQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYS GVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAVTFYLILSTSGIS IETTELEVELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGS AFLNIM
WPHEIANGKWL LYPMQVELEGGQGPQGKGLCS PRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCLY SFDRAAVLHVWGRLWNSTFLEEY
SAVKSLEIVIVRANITVKSSIKNLMLRDASTVIPMVYLDPM AVVAEGVPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPIAADGHP ELGPDGHPGPGTA

Important features:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-
407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408
and 1031-1047

778/237

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATTATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTTCAGATGGTCATACCTACTCTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAAGAAGACAAAACATTGCTGAGGCCTGAGAGAAGCAGATTTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAACAGACTTGATACAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAAATATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTTGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC
CATGATGTATACATTTTGATTGATGACAGTTGAAATCAATAAATTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACCTTTATGTTTTAAAT
AAGAATCATTTGCTTTGAGTTTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTCAGGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAG
ATAATTCTAAGTGAAATTTAAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

779/237

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLRPERSRFDTSILPI
CKDSLGMFNRDLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCEDEDGYKPTQCHGSVGQCWCVDYRGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:**Signal peptide:**

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

780/237

FIGURE 180

CAGACTCCAGATTTCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
CTCTTTAGCCCCGGGATCGCCCCAGCAGGGATGGGCGGACAAGATCTGGCTGCCCTTCCCCGTGCTCTTCTGGCC
GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGGCGCTTCACACCTTCCCTCGATAGCGACTTCACCTTTACCCTT
CCCCCGCGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTTA
GATGGAGCAGGATTAGATATTGATTTCCATCTTGCTCTCCAGAAGGCAAAACCTTAGTTTTGAACAAAGAAAA
TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTCAGCACCATT
TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
AAATATATTACTGGCACAGATATATTGGATATGAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
AGACTAAGCAAAAGTGGGCACATACAAATTCTGCTTAGAGCATTGGAAGCTCGTGATCGAAACATACAAGAAAGC
AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCAATGGTGGTGTGAGCCATTCAAGTTTTAT
ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAGTGAAGCTTAAACTCCAAACTAGAGTACGTAACATTGAAA
AATGAGGCATAAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTGAGATATA
AAAGTAGGAAACAGGTATAATTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
TTGTACTTAAGTGTGTAACAGGAATATTTGCGAATATAGGTTTAACTGAATGAAGCCATATTAATAAGTATTA
TTTCTTAAGTTGAAAAATTTGCAAATGCTTAGCTGAGTTAAATGAATGAGTATTGGGCTAATTGCAACACC
AGTCTGTTTTTAACAGGTTCTATTACCCAGAACTTTTTGTAAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
TCAGTTTTAAGTTATAAATCACCTGAGAATTACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCCAA
CTTTTCTCTATTTACATATGCATCTCTCTATAATGTAATAGAAATAATAGCTTTGAAATACAATTAGGTTTTAT
AGATTTTATAAACCATAATTTTCACTGTAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGTCTACATTC
CCAAAAGCTGACATTTTACAGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
AAATGAAGAATATAGTTTAAAGCTTCTCTCCATAGGGACACATTTCTCTAACCCTTAAGTAAAGTGTAGGA
TTTTAAATTAATGTGAGGTAAAATAAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
TAATCATGTTATGTTAATTTTAACATGTTGCTGAGTTGATAAATCATTATTACCAGGCTAATTGAAGGAATA
TTGCTAAAATGATCTGGGCTACCATAAATAAATATCTCTTTTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
AGAATTTAGAAAACCTTGAAGAAACCTAATCCAAAATAAATTCACTTAAGTAGAAGTATAAATAAATATCTAGA
ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
AACTTTGGCTGTAGGTTTTTATTTTCTACAAGAATTTGTTTGAATTTATTTTGTAGGAGGTACATTTTATA
AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTTATAAAATTTTTATTATAACAACTTTTAT
TAAATGGCCTTTCTGAACACTTTATTTATGTATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
TCAAGTACTAATTTAATTTAATTCATCATGAATGAAGTATAAATTTAAGTTATGCCATTTATAACGTTGTTTTAT
GACTACATTTGTGAGTTAGAAACAACTTAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCAATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
TCTTAGGCCCCCTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAAAATGAGGTAAATGCCGTAT
ATGATCAATTACCTTAAATGGCCAAGAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGCAGAACAA
AGGTCAATAAGATCCTTGCTATGAATACCCCTCCCTTTTGGCGTGTTAAATTTGCAATGAGAAGCAAAATTTACA
GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTCTATAAACTGTGATTAAGAATTCTA
CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT
ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTGAGTAACTTCTTACCATATAAAAACGATAAATTGCTT
TATTTGGAAAAGAAATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAGACAGATATTTAAGAGG
CATAACCAAAAAGCAAACTTGTAACAGAGTAAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
CATATGCTTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATGAGTAAGCTGTTTATCACTT
AACAGCTCATTTTGTCTTTTCAATATACAATTTTAAAAATACTACAATATTTAACTAAGGCCCAACCGATTTC
CATAATGTAGCAGTTACCGTGTTCACCTCACACTAAGGCCTAGAGTTTGTCTGATATGCATTTGGATGATTAAT
GTTATGCTGTTCTTTTCAATGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTATGGTAAATTAATCCTTCTTA
CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATGTATGTCTCCTCAATGAAGATTCTTTAT
GTGAAATTTTAAAGACATTGATTCGCATGTAAGGATTTTTCATCTGAAGTACAATAATGCACAATCAGTGTTG
CTCAAACTGCTTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGATATAATAA
AAATTATCAAAGGAAAA

7P1/237

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPPFVLLLLAALPPVLLPGAAGFTPSLSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFTEL
ILDNMGEQAQEQEDWKKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

782 / 237

FIGURE 182

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACTATGCTGCCTCCCATGGCCCTGCCCAGTGTGTCCTGGATG
CTGCTTTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACTCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACTTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTCATGATCCTCC
TTCTTTTTCCTTTTCTTCACCTTCATTTAGGCTTTTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

783/237

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:**Signal peptide:**

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

784/237

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCTGTCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

785/237

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

7861237

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCAAGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGA
GACACTTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTCTAAAGGTGCTGCACTGCCTATTTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAAAAAAAAAA

787/237

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGS HGLEIFQRCYCGEGLSCRIQ
KDDHHQASNSSRLHTCQRH

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

788/237

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

789/237

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTACCAATCCCCTGCGCCGCGG
CTGGGCCGTCGGAGAGTGCGTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCGCCGCCA
GGGTTTTGAGGATGGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTTGTGGAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTACAGATTAATATTTTTGGGGACAGATTTGTGATGCTTGATTCACCCT
TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATATTACATCAACTGGAACCAGCAGTGAATCTTAATGTTTAC
TTAAATCAGAACTTGCATAAGAAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAAAG
GATCATTCTCTGTTTTCTGATAGTGTATATGGCCATTTTAGTGGGCACAGATCAGGATTTTACAGTTTACTTGG
AGTGTCCAAAATGTCAAGCAGTAGAGAAAATAAGACAAGCTTTCAGAAATTTGGCATTGAAGTTACATCTTGATAA
AAACCCGAATAACCCAAATGCACATGGCGATTTTTTAAAAATAAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTTGGTATTATGATGATGATCCTGAAATCATAACATTGGAAAGAAGAGAATTTGATGC
TGCTGTTAAATTTGGAGAACTGTGGTTTGTAAATTTTTACTCCCCAGGCTGTTACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAGTGGATGGGTACTTTCGAATTGGAGCTGTTAACTGTGGTGATGATAGAAT
GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCCAGTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTTAGTGAGTTTTGCAATGCAGCATGTTAGAAGTACAGTGACAGAATTTG
GACAGGAAATTTTGTCAACTCCATACAAATCTTTTGTGCTTGGTATTTGGCTGGCTGATCCTTTTGTTCATAA
AGGAGGAGATTTGTTTGACTTCACAGACACGACTCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTGGATGCTAA
AGAAATATATTTGGAAGTAATACATAATCTTCCAGATTTTGAAGTACTTTTCGGCAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTTCAATTTGGAAAAATGAAAATCAAATGATCCTGAGCTGAAAAAAT
AAAAATCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTTGGCTGCTTAAGTTTGGTACACTGATCAAGTTGATGA
TCTGTATGTTTTTCAGCCGCTCTAGCAGTATTTAAAGGACAAGGAACCAAGAATATGAAATTCATCATGGAAA
GAAGATTCTATATGATATACTTGCCCTTTGCCAAAGAAAGTGTGAATTCATGTTACCACGCTTGGACCTCAAAA
TTTTCTGCCAATGACAAAGAACCATGGCTTGTGATTTCTTTGCCCCCTGGTGTCCACCATGTCGAGCTTTACT
ACCAGATTCTTACGAAGCATCAAATCTTCTTTTGTGCTTGGTGAAGTTTGGTACACTGATCAACGTTAGTA
GGGACTCTGTAACATGTATAACATTCAAGCTTATCCAACAACAGTGGTATTCACCAGTCCAACATTTCATGAGTA
TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTCATAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACCTTCAACGAAGTGTACACAAAGAAAAACAACGAAGTCTGGATGGTTGATTTCTATTCTCCGTG
GTGTCATCTTGTCCAAAGTCTTAATGCCAGAATGGAAAAGAAATGGCCCGGACATTAACCTGGACTGATCAACGTTGG
CAGTATAGATTGCCAACAGTATCATTCTTTTTGTGCCAGGAAAACGTTCAAAGATACCCTGAGATAAGATTTTT
TCCCCAAAATCAAATAAAGCTTATCAGTATCACAGTTACAATGGTTGGAATAGGGATGCTTATTCCTGAGAAT
CTGGGGTCTAGGATTTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTTCAGTGAAAAAGTTCTACAAG
GAAAAATCATTGGGTGATTGATTTCTATGCTCCTTGGTGGACCTTGCCAGAATTTTGGCTCCAGAATTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAAGTGAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAAGTGTAAAGTTTATTTCTACGAAAGAGCAAAGAGAAATTTTCAAGAAGAGCA
GATAAATACCAGAGATGCAAAAGCAATCGCTGCCTTAATAAGTGAAAAATTTGAAAATCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTTGATAATGTTGAAGATGAAGAAAAAGTTTAAAGAAATTTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACATTTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGTACTGCCA
GAATTATCTACAGCACTGGTGTAAGAAGAGGGTCTGCAAACCTTTTTCTGTAAAGGGCCGGTTTATAAATATTTTA
GACTTTGCAGGCTATAATATATGGTTACACATGAGAACAAGAATAGAGTCATCATGTATTCTTTGTTATTTGCT
TTTAAACACCTTTAAAAAATATTAACAGATTCTTAGCTCAGAGCCATACAAAAGTAGGCTGGATTTCAGTCCATG
GACCATAGATTGCTGTCCCCCTCGACGGACTTATAATGTTTTAGGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGTCTAAGTTGTATAAAGTCCACTTTCCCTTCAGTTTTTTGGCTGACCTGAAAAGAGGTAAC
TAGTTTTTGGTCACTTGTCTCCTAAAAATGCTATCCCTAACCATATATTTATATTTTCGTTTTAAAAACACCCAT
GATGTGGCAGAGTAACAAACCCCTGTTATGCTGATTTATGAGGAGATTCTTCATGTTTTCTTTCTCTCA
AAGGTTGAAAAAATGCTTTTAATTTTTTACAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAGTGCACAAATCTGTAGTTTGTGTATCATCCAGGAAAACCTGAGGGAAAAAATTA
TAGCAATTAAGTGGGCATTGTAGAGTATCCTAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTGTTTCATGTATTTTCTGAAATGCTTTTCATAGAAATTTCCCACTGATAGTTGATTTTTTGAGGCATCTAATAT
TTACATATTTGCTTCTGAACCTTTGTTTTGACCTGTATCCTTTATTTACATTGGGTTTTTCTTCATAGTTTTGG
TTTTTCACTCCTGTCCAGTCTATTTATTATCAATAGGAAAAATTTACTTTACAGGTTGTTTTACTGTAGCTTAT
AATGATACTGTAGTTATTTCCAGTTACTAGTTTACTGTGAGAGGGCTGCCTTTTTTCAGATAAATATTGACATAATA
ACTGAAGTTATTTTTATAAGAAATCAAGTATATAAATCTAGGAAAGGATCTTCTAGTTTCTGTGTTTGTAGA
CTCAAAGAATCACAAATTTGTGAGTAACATGTAGTTGTTTGTAGTTAATTCAGAGTGTACAGAATGGTAAAAAT
CCAATCAGTCAAAGAGGTCAATGAATTAAGGCTTGAACCTTTTTCAAAAAAAAAAAAAAAAAA

790/237

FIGURE 190

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRD LKRIILCFLIVYMAILVGT DQDFYSL LGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKINRAYEVLKDEDLRKKYDKYGEKGLDNQGGQYESWNYRYDFGI
YDDDPEIITLERREFDAAVNSGELWVFNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVN SYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAF AAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICS NLYVFQ P
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTLG PQNF PANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSN IHEYEGHHS
AEQILEFIEDLMNPSVVS LPTTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYS LRIWGLG
FLPQVSTD LTPQTFSEKVLQGNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVKAGKVDC
QAYAQCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAALISEKLET LRNQGKRKDEL
```

Important features:**Endoplasmic reticulum targeting sequence.**

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

797/237

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTCATTCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCACCA
AAAATCCAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAATGAAATGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTTCACATTTTT
TCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCTGTTTTTCCTT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTTATTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTTCAGCATTTGAAAGATTTCCCTAGCCTCTTCCTTTT
TCATTAGCCCCAAAACGGTGCAACTCTATTCTGGACTTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAAGTGGACCCTCTATATTTCCCTCCCTTTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAACTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

792/237

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECKRLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTTFEVNIGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

793/237

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGCACACGGGGCCGCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCC
GAGACTTGAAAGCTTTCGCGTTGCTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

794/237

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPPLPTPGPDRDRELT
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFTPTEKRAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRMIVLS
GSLLRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELNEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETDEDAQQLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:**Signal peptide:**

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

795/237

FIGURE 195

TCGGGCCAGAATTCGGGCACGAGGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTTCGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGGGCTGCGGCAGTGTTCCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGACCCCCGTGGACGCCCCGATATCCCTTCCATTTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

796/237

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS

Important features:**N-glycosylation site.**

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

797/237

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACCTG
CTGTTTCTTCTTACCATTTCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCTGGCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCCGATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTTCAGCCAGGTTTCCTGTGCGCCGCCG
CCTCTGCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

7981237

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRIIPVDLPEARCLCL
GCVNPFTMQEDRSMVSVFVSQVPVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCCGCTTCCTTGAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAAC**ATG**GCGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGCCCCGAGGAGGCCGCGCTGCCGCCGGAGCAGAGCCGGG
TCCAGCCCCATGACCGCCTCCAAC TGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATT CAGAATGGGAGGCTTTTGCAAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAGGTTTGAGTG
GCCGCTTCTTTGTCACCACTCTCCCAGCATTTTTTTCATGCAAAGGATGGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAAC TATTTACAGTGACTCTTGGAATT
CCTGCTTGGTGTCTTATGTGTTTTTCGTCATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTTGTAGATGATGAAGAAGA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACCTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCCAGC
TGACACAGAGGTGGTGGAAAGACTCCTTGAGGCAGCGTAAAAGTCAGCATGCTGACAAGGGAC
TG**TAG**ATTTAATGATGCGTTTTCAAGAATACACACCAAACAATATGTCAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAATCCTTAATTTTTTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCATTTACTTAGGGGT CAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCATCCTTTTCTTGTGTAAAGTATTTAT
TTTTGTCAAATGCGAGGAACATCAGGCACCACAGTGCATGAAAAATCTTTCACAGCTAGAA
ATTGAAAGGGCCTTGGGTATAGAGAGAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTTCCAGCATTTCCACCATGGGCATTCAGGCTCT
CCACACTCTTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTAAAAAAAATGTTTTGTCCAAAATATAGTTGTTGTTGATTTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCCTTCAGTGATGTGCTTTTGGTGAA
GAATTAATGAACTCCAGTACCTGAAAGTGAAAGATTTGATTTTGTTCATCTCTGTAATC
TTCCAAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

200/237

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAATAGPEEEAALPPEQSRVQPMTASNWTLVMEGEWMLKIFYAP
WCPSCQQT DSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISSECFYVPLPRHLSESEQNRRSEEAHRAEQLQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEDEEEEDNLAAGVDEERSEANDQGPPGEDGV TRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSQHADKGL

Important features:**Signal peptide:**

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

207/237

FIGURE 201

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTTCAGGGGCTCTCAAGAACAATGGAATATCATCCTGATTTAGAAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTTCAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCAAATTGTG
TATGGATTACGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTCAATGTAAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCCGCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCACTTCTGCCATTGACTTGGCATTACCTTGGGTAAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAACTGAGGGTGAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAATCCTA
CTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAATCTTT
TTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACGG
CAGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACACCCGGCTAATTTTGTATTTTTTTTAGT
AGAGACAGGGTTTTCTCCATGTCGGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGATAATATGTAATTGTAGGGAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTTAAATACA
AAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTT
ACCATTTTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTTGTCTTATTTTTTTTCTACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATA
TTTCTGTATTACTAGGGAGGCATTTACAGTCCTCTAATGTTGATTAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTATGCAGTTTAAATGGATGATTTTATGTTATGTGGATTTTCAT
TTCAATAAAAAAAACTCTTATCAAAAAA

202/237

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

203/237

FIGURE 203

GGAAGGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCAGACCTGGAGGGTCTCGCTCTGTCA
CACAGGCTGGAGTGCAGTGGTGTGATCTTGGCTCATCGTAACCTCCACCTCCCGGGTTCAAGTGATTCTCATGCC
TCAGCCTCCCGAGTAGCTGGGATTACAGGTGGTGAATCCAAAGAGTGAATCCGTCGGAGGAAAATGACTCCCCAG
TCGCTGCTGCAGACGACACTGTTCTGCTGAGTCTGCTCTTCTGGTCCAAGGTGCCACGGCAGGGGCCACAGG
GAAGACTTTTCGCTTCTGCAGCCAGCGGAACAGACACACAGGAGCAGCCTCCACTACAAACCCACACCAGACCTG
CGCATCTCCATCGAAGCTCCGAAGAGGCCCTCACAGTCCATGCCCCCTTCCCTGCAGCCCACCTTGCTTCCCGA
TCCTTCCCTGACCCAGGGGCTCTACCCTTCTGCCTCTACTGGAACCGACATGCTGGAGATTACATCTTCTC
TATGGCAAGCGTGAATCTTCTGCTGAGTGACAAAGCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGAGAGCCTG
GCTCAGGGCCCCCGCTGTTAGCCACTTCTGTCACTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCC
GCCAGCTTACCTTCTCTTCCACAGTCTCCCCACAGGCCGCTCACAATGCCTCGGTGGACATGTGCGAGCTC
AAAAGGGACCTCCAGCTGCTCAGCCAGTTCTTGGTGAAGCTCCCGAGAAGGCCCTCAAGGAGGCCCTCGGCTGCCCTC
GCCAGCCAGCAGTTGCAGAGCCTGGAGTCGAACTGACCTCTGTGAGATTTCATGGGGGACATGGTGTCTTTCGAG
GAGGACCGGATCAACGCCACGGTGTGGAAGTCCAGCCACAGCCGCTCCAGGACCTGCACATCCACTCCCGG
CAGGAGGAGGAGCAGAGCGAGATCATGGAGTACTCGGTGCTGCTGCCCTCGAACACTCTTCCAGAGGACGAAAGG
CGGAGCGGGGAGGCTGAGAAGAGACTCTCTGGTGGACTCCAGCAGCCAGCCCTGTTCCAGGACAAGAATTCC
AGCCAAGTCTGGGTGAGAAGGTCTTGGGGATTGTGGTACAGAACACCAAGTAGCCAACTCAGGAGCCGCTG
GTGCTCACTTTCAGCAGCAGCTACAGCCGAAAGTGTGACTCTGCAATGTGTGTTCTGGGTGAAGACCCACA
TTGAGCAGCCCGGGCATTGGAGCAGTGTGGGTGTGAGACCGTCAGGAGAGAAACCCAAACATCCTGCTTCTGC
AACCATTGACCTACTTTCAGAGTGTGATGCTTCTGGTGGAGGTGAGCGCGTGCACAAGCACTACCTGAGC
CTCCTCTCTACGTGGGCTGTGTCTCTCTGCCCTGGCCTGCCTTGTCACTTGGCGCTTACCTCTGCTCCAGG
GTGCCCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATCAAGGTGCACATGAACCTGCTGCTGGCCGTCTT
CTGCTGGACACGAGCTTCTGCTCAGCGAGCCGGTGGCCCTGACAGGCTCTGAGGCTGGCTGCCGAGCCAGTGCC
ATCTTCTGCTGCTTCTCCCTGCTCAGCTGCCCTTCTGGTGGGCTCGAGGGGTACAACCTCTACCGACTCGTG
GTGGAGGTCTTTCGACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTTGGGGCTTCCCCATCTT
CTGGTGACGCTGGTGGCCCTGGTGGATGTGGACAATATGGCCCATCATCTTGGCTGTGCATAGGACTCCAGAG
GGCGTCATCTACCTTCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATACCAACCTGGGCTCTTCCAGC
CTGGTGTCTTCTGTCAACATGGCCATGCTAGCCACCATGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAA
AAGTGGTCACATGTGCTGACACTGCTGGGCTCAGCCTGGTGGTGCAGATCCTGCGGCTGCGCCCCACACCCAA
TTTGTCTTGGCACCTTCCAGCTTGTGCTCTTCTTCTTCTAGCATCATCACCTCTTCCAAGGCTTCTCTCATC
TTCATCTGGTACTGGTCCATGCGGCTGCAGGCCCCGGGTGGCCCCCTCCCTCTGAAGAGCAACTCAGACAGCGCC
AGGCTCCCCATCAGCTCGGGCAGCACCTCGTCAGCCGCTCTAGGCTCCAGCCACCTGCCATGTGATGAAG
CAGAGATGCGGCTCGTGCACACTGCCTGTGGCCCCGAGCCAGGCCAGCCCCAGGCCAGTCAAGCCGAGACT
TTGGAAGCCCAACGACCATGGAGAGATGGGCCGTTGCCATGGTGGACGGAATCCCGGGCTGGGCTTTTGAATTG
GCCTTGGGACTACTCGGCTCTCACTCAGCTCCACGGGACTCAGAAGTGCGCGGCCATGCTGCCTAGGGTACTG
TCCCCACATCTGTCCCAACCCAGCTGGAGGCTGGTCTCTCTTACAACCCCTGGGCCCAGCCCTCATTGCTGGG
GGCCAGGCTTGGATCTTGGAGGTCTGGCACATCCTTAATCCTGTGCCCTGCTGGGACAGAATGTGGCTCCA
GTTGCTCTGTCTCTCGTGGTCACTTGGGCACTCTGCATCCTCTGTCTTCTTAACTCAGGTGGCAGCCAGGG
CGAATGGGGCCAGGGCAGACCTTCAAGGGCCAGAGCCCTGGCGGAGGAGAGGCCCTTGGCAGGAGCACAGCAGC
AGCTCGCTACCTCTGAGCCAGGCCCTTCCCTCCCTCAGCCCCCAGTCCCTCCCTCCATCTTCCCTGGGGTTC
TCCTCTCTCCAGGGCTCCTTGTCTCTTCTTCTCAGCTGGGGGTCCCCGATTCCAATGCTGTTTTTGGGGA
GTGGTTTCCAGGAGCTGCCTGGTGTCTGCTGTAAATGTTTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCA
GGCTCGGTACCGATGCGTGGGCTGGGCTAGGTCCCTCTGTCCATCTGGGCTTGTATGAGCTGCATTGCCCTTG
CTCACCTGACCAAGCACAGCCTCAGAGGGGCCCTCAGCCTCTCTGAAGCCCTCTTGTGGCAAGAACTGTGGA
CCATGCCAGTCCCGTCTGGTTTCCATCCCACTCCAAAGGCTGAGACTGACCTCCTCTGGTGACACTGGCCTA
GAGCTGACACTCTCTAAGAGGTCTCTTCAAGCCCCCAATAGCTCCAGGCGCCCTCGGCCGCCATCATGGT
TAATTCTGTCCAACAAACACACAGGGTAGATTGCTGCCTGTTGTAGGTGGTAGGGACACAGATGACCGACCTG
GTCACCTCCTGCTGCCATCCTGGAAGACTTCTCTCAGGAGTCAAGGTTCAATCTTACCTTGAAGAT
GGGAAGGATGTTCTTTTACGTACCAATTCTTTTGTCTTTTGTATTTAAAAAGAGTACATGTTTATTGTAGAGA
ATTTGGAACTGTAGAAGAGAATCAAGAAGAAAAATAAAATCAGCTGTTGTAATCGCCTAGCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

204/237

FIGURE 204

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921

<subunit 1 of 1, 693 aa, 1 stop

<MW: 77738, pI: 8.87, NX(S/T): 7

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRCSQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRD FLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLT FQHQLQPKNVT LQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSA LACLVTIAAYLCSRVP LPC
RRKPRDYTIKVHNMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVT LVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNMA MLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGF LIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI

Important features:**Signal peptide:**

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

205/237

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCCTGGATGGGCCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACCTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTCTTCTGTTCAACATGG

206/237

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTGGCTTTGA
TCCTTTTCAAAAACCTGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTTTGGATGGGATTATGTGGAACTACCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCCACTGCAGCCTTCCCCTGGCGGTGGTGAAAGAGAC
TCGGGAGTCGCTGCTTCCAAAGTGCCCGCCGTGAGTGAGCTCTCACCCCACTCAGCCAAATGAGCCTCTCGGGC
TTCTCCTGCTGACATCTGCCCTGGCCGGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTCC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTCTCTACTATCTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAAATGTATGGATACAACCTTACGTTTGTAGAAAGATTTGGGCTTGAAGACCCAGAAGATGACATATGCAAGT
ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGGTACTGTACCAG
GAAAACAGATTTCTAAAGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTCTTCTGAACCAGGGT
TCTGCATCCACTACAACATTGTCTATGCCACAATTACACAGAAGCTGTGAGTCTTTCAGTGCTACCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGAAGACCTTATTCGATATCTTGAACCAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTGTTTTGGAA
GAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCT
CAGTGTCCATAAGGGAAGAACTAAAGAGAACCAGTACCAATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGT
GTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
ACGAGGTCTTTCAGTTGAGACCAAAGACCGGTGTGAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGTCAGAGGGAGCACAGGAGGATAGCCGCATCACCCAGCAGCTCTTGCCCA
GAGTGTGTCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGT
TCAAGGACCTTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGAGTTGTGCA
ACAGCTCTTTTGAAGGAGGCGCTAAAGGACAGGAGAAAAGGCTTCAATCGTGGAAAGAAAATTAATGTTGTAT
TAAATAGATCACAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTTCAGTTCTTTC
GATACGGCTTAGGGTAATGTCTAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAAC
TCTAAAGCTCCATGTCTGGGCTTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTTTTTGTCTCATATTCACAT
ATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAACCTGT
GTCATGCTGATAGGACAGACTGGATTTTTTCATATTTCTTATTAATAATTTCTGCCATTTAGAAGAAGAGAACTACA
TTCATGGTTTGGAAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTG
TTTCATTGTGTACATTTTTATATCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCT
ATTTTTACCAAAGGTATTTAATATCTTTTTTATGACAACTTAGATCAACTATTTTTCAGCTTGGTAAATTTTCT
AAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAA
GACTTTTTGAAAATAAATTAAATTATCATATCTTCCATTCTCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGA
AAGTAGACATTGAGATCCAGCCATTACTAACCTATTCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
AAAGCACCTTGAAAAAGACTTGGCAGCTTCTGTATAAGCGTGCTGTGCTGTGCTGTCAGTAGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTTTATTATCTTAAACTCTGTTCCATACACTGTATATAATACATGGATATTTTTATGTACA
GAAGTATGTCTCTTAACAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAAATATTTTGGTTGT
AAAATGCTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGATTGAATCATCAATAAAGA
ATGTGGCTATTTTGGGGAGAAAATTAAAAAAAAAAAAAAAAAAAGGTTAGGGATAACAGGGTAATGCGGCC

207/237

FIGURE 207

MSLFGLLLLT SALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSPALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKRKSRVVDLNLLEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

208/237

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCCATTTTGGGAAGA
AGACTAAAAATGGTGTTTCCAATGTGGACACTGAAGAGACAAATTCTTATCCTTTTAAACATAATCCTAATTTCC
AAACTCCTTGGGGCTAGATGGTTTCTAAAACTCTGCCCTGTGATGTCACTCTGGATGTTTCAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTTGACAGAAATTCCTGGAGGTATTCCCACGAACACCACGAACCTCACCTC
ACCATTAAACCACATACCAGACATCTCCCCAGCGTCCTTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCAACTGTGTACCTATTTCACTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCCAGAAGC
TTTAGTGGACTCACTTATTTAAAAATCCCTTTACCTGGATGGAAACCAGCTACTAGAGATACCGCAGGGCCCTCCCG
CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACAACTCTTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTCTACCTGGGCCAAAACCTGTTATTATCGAAATCCTTGTTATGTTTCATATTTCAATAGAGAAA
GATGCCTTCTAAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAATGTACAGCCGTCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACACATGATTGCAAAAATCCAAGAAGATGATTTTAATAAC
CTCAACCAATTACAAATCTTGACCTAAGTGGAAATGCCCCTCGTTGTTATAATGCCCCATTTCTTTGTGCGCCG
TGTAATAATAATTCTCCCTACAGATCCCTGTAAATGCTTTTGATGCGCTGACAGAATTAAGTTTACGTCTA
CACAGTAACTCTTTCAGCATGTGCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC
CAAACTTCTTGGCCAAAGAAATGGGGATGCTAAATTTCTGCAATTTCTCCCGAGCTATCCCAATGGATGTG
TCTTTCAATTTTGAACCTCAGGTCTATCGTGCATCTATGAATCTATCACAAGCATTTTCTTCACTGAAAAGCCTG
AAAATCTGCGGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAACCTCTCGCCATTACATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAATTGCTAACCTCAGCATGTTTAAACAATTTAAAGA
CTGAAAGTCTAGGAAATCTTTCAGTGAATAAAATATCACCTCAGGAGATTCAAGTGAAGTTGGCTTCTGCTCAAAC
GCCAGAACTTCTGTAGAAAGTTATGAACCCAGGTCTGGAACAATTACATTATTTAGATATGATAAGTATGCA
AGGAGTTGCAGATTCAAAAACAAAGAGGCTTCTTTCATGTCTGTTAATGAAAGCTGCTACAAGTATGGGCAGACC
TTGGATCTAAGTAAAAATAGTATATTTTTTGTCAAGTCTCTGATTTTCAGCATCTTTCTTCTCAAATGCCTG
AATCTGTGAGGAAATCTCATTAGCCAACTCTTAATGGCAGTGAATTTCAACCTTTAGCAGAGTGAAGTATTTG
GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCATTTGAAGAGCTTCACAACTGGAAGTTCTGGAT
ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTACTCATATGCTAAACTTTACCAAGAACCTAAAGGTT
CTGCAGAACTGATGATGAACGACAATGACATCTCTTCTCCACCAGCAGGACCATGGAGAGTGAGTCTCTTAGA
ACTCTGGAATTCAGAGGAAATCACTTAGATGTTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGCCCTTCTGGAGTTTTTGATGGT
ATGCCCTCAAATCTAAAGAATCTCTCTTTGGCCAAAATGGGCTCAAATCTTTCAGTTGGAAGAACTCCAGTGT
CTAAAGAACCTGGAACTTTGGACCTCAGCCACAACCAACTGACCACTGTCCCTGAGAGATTATCCAACCTGTTCC
AGAAGCCTCAAGAATCTGATTCTTAAGAATAATCAAACTCAGGAGTCTGACGAAGTATTTTCTACAAGATGCCTTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCTC
ACAATCTGAAGATGTTGCTTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACCATACGGAGGTGACTATTCCTTACCTGGCCACAGATGTGACTTGTGTGGGGCCAGGAGCACACAAGGGC
CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTCACTTTCCATA
TCTGTATCTCTCTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTTCTGGGATGTGTGGTATATTTACCAT
TTCTGTAAGGCCAAGATAAAGGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGTGGCCAACTGGAAGACCCAAGAGAGAAA
CATTTTAATTTATGTCTCGAGGAAAGGGACTGGTTACCAGGGCAGCCAGTTCTGGAAAACCTTTCCAGAGCATA
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTTAAGATAGCATTTTAC
TTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGTGATTATCTTGATATTTCTTGAGAAGCCCTTTCAGAAG
TCCAAGTTCTCCAGCTCCGAAAAGGCTCTGTGGGAGTTCTGTCTTGAAGTGGCCAAACAAACCCGCAAGCTCAC
CCATACTTCTGGCAGTGTCTAAAGAACGCCCTGGCCACAGACAATCATGTGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTTGCAAAACACAACCTGCCTAGTTTACCAAGGAGAGGCTGGC

209/237

FIGURE 209

MVFPMWTLKRQILILFNIIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTTLTINHIPDISPASFHRLDHLVEIDFRNCNCPVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLQNCYYR
NPCYVSYSIEKDAFLNLTKLKVLSLKDNNTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSLKSL
KILRIRGYVFKEKLSFNLSPLHNLQNLVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGO
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVLOKLMMNDNDISSSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSLFLPSGVFDGMPPNLKNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLO
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTDAVWFVWWVNHTVETIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLISISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIIILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAACATGGAAAAACATGTTCTTC
AGTCGTCAATGCTGACCTGCATTTTCCTGCTAATATCTGGTTCCCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGGTGGGCAAAATATGTGACAGAACTAGACCTTGTCTGATAATTTTCATCACACACATAACGAATGAAT
CATTTCAAGGGCTGCAAAATCTCACTAAAATAAATCTAAACCACAACCCCAATGTACAGCACCAGAACGGAAATC
CCGGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCTCAACCTAAAAAACCTAAGGGAGTTACTGC
TTGAAGACAACCAAGTTACCCCAATACCCCTCTGGTTTGGCAGAGTCTTTGACAGAACTTAGTCTAATTCAAAACA
ATATATAACAATAAACTAAAGAGGGCATTTCAAGACTTATAAACTTGAAAAATCTCTATTTGGCCTTGAAGCTGCT
ATTTTAACAAAGTTTGGGAGAAAACTAACATAGAAGATGGAGTATTTGAAACGCTGACAAATTTGGAGTTGCTAT
CACTATCTTTCAATTTCTTTTCACACGTGCCACCCAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACA
CCAGATCAAATACATTAGTGAAGAAGATTTCAAGGGATTGATAAATTTAACATTACTAGATTTAAGCGGGAAC
GTCCGAGGTGCTCAATGCCCATTTCCATGCGCTGACCTTGTGATGGTGGTGCTTCAATTTAGATCGTGTCTT
CTTTTCAAACTTGACCCAACTTCGATACCTAAACCTCTCTAGCACTTCCCTCAGGAAGATTAATGCTGCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTTGAATTCACATATTTAGTGGGAGAAATAGTCTCTGGGGCAT
TTTTAACGATGCTGCCCGCTTAGAAATACTTGACTTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCATA
TTAATATTTCCAGAACTTCTCTAACTTTGTCTCAGCGGCATTGCATTTAAGAGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCTGATGCAGCTTCCAACTTATCGACTATCAACTGGGTATTAATTTTATTA
AGCAAATCGATTTCAAACCTTTTCCAAAATTTCTCCAATCTGGAAATTAATTTACTTGTGAGAAAACAGAAATATCAC
CGTTGGTAAAGATACCCGGCAGAGTTATGCAAAATAGTTCCTCTTTCAACGTCATATCCGGAAACGACGCTCAA
CAGATTTTGAGTTTGACCCACATTCGAACCTTTATCATTTCACCCGTCCTTTAATAAGCCACAATGTGCTGGT
ATGGAAGAGCTTAGATTTAAGCCTCAACAGTATTTTCTTCAATTTGGGCCAAACCAATTTGAAAATCTTCTGACA
TTGCTGTTTAAATCTGTCTGCAATAGCAATGCTCAAGTGTAAAGTGGAACTGAATTTTCAGCCATTCTCTCATG
TCAAATATTTGGATTTGACAAACAATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATGTCCGACTTGG
AAGTATTTCCAGACTTCACTATAATTCACACTATTTGACTAGTCAAGGCGTAACACATCATCTAGATTTTCAAG
ATTTACAAATCTAAAAGTTTAACTTGAGCCACAACAACATTTATACTTTAACAGATAAGTATAACCTGGAAA
GCAAGTCCCTGGTAGAATTAGTTTTAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAACAGGTATA
TCTCCATTTTCAAAGGTCTCAAGAATCTGACACGTCTGGATTTATCCCTTAATAGGCTGAAGCACATCCCAATG
AACTAAGACCACCACCAAAATTTATCTATGTTGGAATCAACGGAACCCCTTTGAATGCACCTGTGACATTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAATTTCCAGACTGGTAGATGTCATTTGTGCCAGTCTG
GGGATCAAAGAGGGGAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTTCAGATGTCAGTGCAGTGATATTAT
TTTTCTTACGTTCTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTCACCATTGTGTTTACTGGGATGTTT
GGTTTATATATGTTGTTAGCTAAGGTAAAGGCTACAGGTCTCTTTCCACATCCCAATTTTAACTGATG
CTTACATTTCTTATGACACCAAGATGCCTCTGTTACTGACTGGGTGATAAATGAGCTGCGCTACCACCTGAAG
AGAGCCGAGACAAAAACGTTCTCTTGTCTAGAGGAGAGGGATTGGGACCCGGGATTGGCCATCATCGACAACC
TCATGCAGAGCATCAACCAAGCAAGAAAAACAGTATTTGTTTTAACCAAAAAATATGCAAAAAGCTGGAACTTTA
AAACAGCTTTTACTTGGCTTTGACAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCTTGTGGAGC
CAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCTCCAGTGGCCTGACA
ACCCGAAGGCAGAAAGGCTTGTGTTGGCAAACTCTGAGAAATGTGGTCTTGACTGAAAATGATTACGGTATAACA
ATATGTATGTCGATTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTCCGCGCCATAATAAAGATGCAAG
GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAACAAATTTATCCCAAACTTAGTGGTTTAAACACACA
TTTGTGCGCCACAGTTTTTGGGGTCCAGGAGTCCAGGCCAGCATAACTGGGTCTCTGCTCAGGGTGTCTCAG
AGGCTGCAATGTAGGTGTTTACCAGAGACATAGGCATCACTGGGGTCACACTCATGTGGTTGTTTTCTGGATTCA
ATTCCTCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCCTCTCCACAAAGGCAGCTTGTCTC
ATCAGAGCTAGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTTTGAATCCAATCAAAAAAGTGAT
ATCTCATCACTTTGGCCATATTCTATTTGTTAGAAAGTAAACCACAGGTCCCAACAGCTCCATGGGAGTGACCACC
TCAGTCCAGGGAAAAACAGCTGAAGACCAAGATGGTGAGCTCTGATTGCTTCACTTGGTTCATCAACTATTTCCCT
TGACTGCTGTCTGGGATGGCTGCTATCTTGATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTGGCACTGTGACTAATGTGCTCTAATA
TTAAGCTGTTGTTTATATTTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCGTTCCGGTTTTAT
TTACAGTTGCTTTTACAAATATTTGCTGTAACATTTGACTTCTAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTAAAAAGTATGCAGCTAAATTCGAAGCTTTTGGTCTATA
TTGTTAATTGCCATTGCTGTAATCTTAAATGAATGAATAAAAAATGTTTCATTTTACAAAAA

271/237

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSPCDEKKQNDSVIAECSNRRRLQEVPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHONGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTLRLYNLSSTSLRKINAAWFKNM
PHLKVLDLEFNVLVGEIVSGAFLTMLPRLEILDLSFNLIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFEFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIPHVKYLDLTNNRLDFDNASALTELSDELEV
LDLSYNSHYFRIAGVTHHLEFIQNFTNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDLNRLKHIPNEAFLNLPASLTEHINDNMLKFFNWT
LLQQFPRLLELLDLRGNKLLFLTDSLSDFTSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDL
SNLLKTINKSALETKTTTKLSMLELHGPNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITTMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTFFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLAIIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTA FYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLEFWQTLRNVVLTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAAC TGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCCGTGTCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCCGGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTCCTTCCTG
GAGGAGCAGCTGGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTG
GACTGAGCCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACAATAAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGCCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

273/237

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

214/237

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGGGCCCCGGGGAGGGGAACTGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCCCTGTCCGGGGGATGACTGATTCT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGGCCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCCCTGTCTCCGAGTCGTTG
TGCAGCGTGTGTACCAGCCCTTCCTCACACCTGCGACGGGCACCGGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTA
CGCGTGCTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCCTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCCTCCTCCCTTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGAC
CCCCAGCACAATAAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

216/237

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCCTGTCTCCGAGTCGTTTCGTGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGAC
GGGCACCGGGCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGCAGCCCTGG
GCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTTGGGAGGGGACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGCCCCACTGCACA
GCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCCTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTCCTGGAGGAGCAGCT
GGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCCTCCTCCCCCTTCCTCGGGAG
GCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGG
CCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDVPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSL SADGTLCV PKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

218/237

FIGURE 218

GGTTGCCACAGCTGGTTTATGGGCCCGACCACTGGGGCCCTTGTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCTGGGACGGTCCGTTTCTGCGTCAGCTGCCGGCCG
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGCTCTCCCTTCTCCGCTGGGCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTCTCGTGTAATCGCAAAACCCATTTTGGAGCAGGAATTCCAATCA
TGTCTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGGAACACCTTTTGCTGTG
ATGGCCCGCTCATGATGGCCCGGCAAAAGGGCATTCTTCTACCTGACCTTTTCTCATCTGGGGACATGTACAC
TCTTCTTCGCTTTTGAGTGCCGCTACCTGGCTGTTTCAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCT
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTCAGTGACCTGGAGTGATTCTCGGGCGCTACCAGATG
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCCTCGTA
TCAAGAAATTTCCAGATAAAACAACAGATTGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGCCTCCCATTTGCAGCATCTGTGACAACGTGTGGAGCGCTTCGACCATCACTGCCCCTGGGTGGGCAATTGTG
TTGGAAGAGGAACACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCCTCCTCACAATCTATGTCTTCGCCT
TCAACATCGTCTATGTGGCCTCAAATCTTTGAAAATTGGCTTCTTGGAGACATTGAAAGAAACTCCTGGAACCTG
TTCTAGAAGTCTCTATTGCTTCTTTACACTCTGGTCCGTGCTGGGACTGACTGGATTTTCACTTTTCTCGTGG
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGGCCCCCTTGCCCCCAGTGCTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGAAGTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC
CAGCCCCACAGAACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGGCTATGAGAGATTTCAAGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAAC
GTTTTTCTTTGGTCTTTAGTCACCCAGTTGCACACTGGCATTCTTCTGCTGCAAGCTTTTTTAAATTTCTGAAC
CAAGGCAGTGGCAGAAGATGTAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCCTTGGACCCCTCTCTTCCCTCCAGATCCCAGCCCTCCTGCTTGGGGTCAC
TGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCCACAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNTTGTCTTCTTTGACTCCTGCTCCCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCTCAGAGGAAGCCCCGAGTGCTCACTTAAAC
ACTATCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG
GCTCTCCTCTCCTCTCCTCTCCCCCGATGTACCCTCAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAGCCT
CGGCTGAGTGAGGGAAAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCCACCTCTGGCT
ATGGTAACCACACTGGGGGCTTCTCCAAGCCCCGCTCTTCCAGCACTTCCACGGCAGAGTCCCAGAGCCACTT
CACCTGGGGGTGGGCTGTGGCCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTAGGGAAGAAGATTTATGT
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT
CGGGTGGGGGAGTGTAACCGGAACTTTTCATCTATTTGAAGGCGATTAAACTGTGTCTAATGCA

279/237

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGFYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNRYFYFYL
FILSLSLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLCFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGFLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCTTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGACAAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATTTCT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACCTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCGTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTGAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAI FTGLAALCLFQGV PVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNV DVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLT CIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPVVRVKVT VNYPPYISEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKG VKVENRPFLSKLI FFNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

224/237

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAGGGGACCGGGGCCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
ATCACCGCCTGGCCCCGACTCCACCAATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCACAGCACCTGCCTTACA
GAGGCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTAC
CAGTTCTCCTGTGGGGCTGGATTTCGGAGGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCACGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC
AATGTTATCCAGGTGGACCACTGCTGGGCTCTTCTGCCCTCTCGGGATTACTACTTAAACAGAAGTGCCAATGAG
AAAGTGCACAGCTATCTGGATTACATGGAGGAAGTGGGATGCTGCTGGGTGGGCGGCCACGCCACTGAGAG
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGGCGCGAC
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG
CAGGTGTGAGAGCTCATCAACCGCACGGAAACCAAGCATCCTGAACAATTACCTGATCTGGAACTGGTGCAAAAG
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTCCAAACATGTTGAATTTGTACAACCTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG
CCTCCCAGCCGAGACAGTGGAGCATGACCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAATGAGATC
GTCTTCCCCGCTGGCATCCTGCAGGCCCTTCTATGCCCGCAACCAACCCCAAGGCCCTGAACCTTCGGTGGCATC
GGTGTGGTTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCAACGGCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGCGCCGCCAGACGCTGGGGGAGAAACATTACTGACAACGGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
AACCACAGCTCTTCTTCGTGGGATTTGCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTACGAGGGG
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGG
CACTTCGGCTGCCCTGTGGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG
TGCTCTGCTTTGGGGGTGCCCTGCCTCCAGCAGAGCCCCACCATTCACTGTGACATCTTCCGTGTCACCT
GCCTGGAAGAGGTCTGGGTGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESIDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQONQA
ILKHLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTA
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVDYDGYEISEDSF
FQNMLNLNFSKVMADQLRKPPSRDQWSMTPQTVNAYYLPKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHEALTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPGLCEVW

Type II Transmembrane domain:

amino acids 32-57

226/237

FIGURE 226

GCCCCGCCCTCCGCCCTCCGCACTCCCGCCTCCCTCCCTCCGCCCGCTCCCGCGCCTCCTCCCTCCCTCCTCCC
CAGCTGTCCCGTTCCGCGTCATGCCGAGCCTCCCGCCCCCGCGGCCCGCTGCTGCTCCTCGGGTCTGCTGCTGCT
CGGCTCCCGGCCCGGCCCGCGCGGCCCGGCCAGAGCCCCCGTGCTGCCCATCCGTTCTGAGAAGGAGCCGCTGCC
CGTTCGGGAGCGCGCAGGTAGGTGGGCGCCCGGGGAGGCGCGGGCGGGAGTCCGGCTCGGGGCGAGTCAGCGC
CAGCCCCGAGGGGGCGCGGGGCGCAGGTGGCTCGCGCGCGCGGGCGGCCCGGAGGGTGGGCGGGGCGAGAAGGGC
GCGGTGCCCTGGGACCCGGGACCCGCGGGCAGCCCCCGGGGCGGCACACGGCGCGAGCTGGGACAGCGGCCTCCAGC
CAAGCCCCGTCCCCGAGGCTGCACCTTCGGCGGGAAAGTCTATGCCCTGGACGAGACGTGGCACC CGGACCTAGG
GGAGCCATTCCGGGTGATGCGCTGCGTGCTGTGCGCTGCGAGGCGCAGTGGGGTCCGCGTACCAGGGGGCCTGG
CAGGGTCAGCTGCAAGAACATCAACCAGAGTGCCCAACCCCGGCCCTGTGGGACAGCGCGCCAGCTGCCGGGACA
CTGCTGCCAGACCTGCCCCAGGACTTCGTGGCGCTGCTGACAGGGCGGAGGTGCGAGGCGGTGGCACGAGCCCCG
AGTCTCGCTGCTGCGCTCTAGCCTCCGCTTCTCTATCTCTACAGGCGGCTGGACCGCCCTACCAGGATCCGCTT
CTCAGACTCCAATGGCAGTGTCTGTTTGGACACCTGCAGCCCCACCCAAGATGGCCTGGTCTGTGGGGTGTG
GCGGGCAGTCCGTTGTCTCTGCGGCTCCTTAGGCGAGCAACAGCTGCATGTGGCATGTGGCATGTGGCAT
CCCTTCAGGGGAGGTCTGGGGGCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCATCCTGAC
TCTAGAAGCCCCACCAGCAGGGCGTAGGGGGCATCACCTGCTCACTCTCAGTGACACAGAGGACTCCTTGCA
TTTTTTGCTGCTCTTCCGAGGCTTGCAGGACTAACCCAGGTTCCCTTGAGGCTCCAGATTCTACACCAGGGGCA
GCTACTGCGAGAACTTCAGGCCAATGTCTCAGCCAGGAACCCAGGCTTGTGAGGTGCTGCCCAACCTGACAGT
CCAGGAGATGGACTGGCTGGTGTGCTGGGGGAGCTGCAGATGGCCCTGGAGTGGGACAGGCGAGGGCTGCCAT
CAGTGGACACATTGCTGCCAGGAAGAGCTGCGAGCTCCTGCAAAGTGTCTTTGTGGGGCTAATGCCCTGATCCC
AGTCCAAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCTGATCCTCCAGGTGCAATT
GGTAGGACAACAGTGAAGTGGTGGCCATGACACTGGGAACCAAGCCTCAGCGGAGGAGTGGCCACTGTGCTCCT
GTGCCACATGGCTGGCCTATCCTCCCCTGCCCCAGGCGCTGGGTATCTGCCCTGGGCTGGGGTGCCCGAGGGGC
TCATATGCTGCTGCAGAAATGAGCTCTTCTGAAAGTGGGACCAAGGACTTCCAGACGGAGAGCTTCGGGGGCA
ACGTGGCTGCCCTGCCCTACTGTGGGGCATAGCGCCCCGCCCTGCCCGTGCCCCTAGCAGGAGCCCTGGTGCTACC
CCCTGTGAAGAGCAAGCAGCAGGGCACGCGCTGGCTTTCCTTGGATACCCACTGTCACTTGCACATGAAGTGCT
GCTGGCTGGGCTTGGTGGCTCAGAACAAGGCACTGTCACTGCCCACCTCCTTGGGCCCTCTGGAACGCCAGGGCC
TCGGCGGCTGCTGAAGGGATTCTATGGCTCAGAGGCGCCAGGGTGTGGTGAAGGACCTGGAGCCGGAACCTGCTGCG
GCACCTGGCAAAAGGCATGGCTTCCCTGATGATCACCAACAGGTAGCCCCAGAGGGGAGCTCCGAGGGCAGCCT
CTCCTCCCAGGTGCATAGCCAACCAATGTGAGGTTGGCGGACTGCGCCTGGAGCGCGCCGGGCGCAGGGGGT
GCGGGCGCTGGGGGCTCCGATACAGCCTCTGCTGCGCGCCTGTGGTGCCTGGTCTCCCGGCCCTAGCGCCCGC
CAAACCTGGTGGTCTTGGGCGGGCCCCGAGACCCCAACACATGCTTCTCGAGGGGCGAGCAGCGCCCCACGGGGC
TCTGTGGGCGCCCACTACGACCCGCTCTGCTCACTCTGCACCTGCCAGAGACGAACGGTGTCTGTGACCCGGT
GGTGTGCCCCACCGCCAGCTGCCACACCCGCTGAGGTTGGCGGACTGCGCCTGGAGCGCGCCGGGCGCAGGGGGT
TTTTGATGGTGACCGAGCTGGCGGGCAGCGGGTACGCGGTGGCACCCCGTTGTGCCCCCTTTGGCTTAATTAA
GTGTGCTGTCTGCACCTGCAAGCAGGGGGGCACTGGAGAGGTGCACTGTGAGAAGGTGCACTGTCCCCGGCTGGC
CTGTGCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCAACAGTGTCCAGGTGAGGCCCCACCCAGCTGGG
GGACCCCATGCAGGTGATGGGCCCCGGGGCTGCGTTTTTGTGGGAGTGGTTCAGAGACAGAGCTCCACAAGGGT
CCCCCTCAGTGCCCCCGTTTGGAGAGATGAGCTGTATCACCTGCAGATGTGGGGTAAGTGGGGAGCAGAGGCTTGT
GTGAGGTGGGTACTGGGAGCCTGGTCTGGAGTAGGGAGACCTTCCAGGGAGGTCCCTGAAGAAGCTGAAGGTCA
CTGTGTCCCAGTGCCCTCTGGGGGACACTCAGTGTCTGCTCTGTCTTGTACCAGGCAGGGGTGCCCTCACTGTGAGC
GGGATGACTGTTCACTGCCACTGTCTGTGGCTCGGGGAAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCCCACC
GGCGGCGTAAGTGAGGGAGTCCAGGCTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCTGTGGGACTCCTGATCAG
GGAAGGGAGCACTCACTGTGTGCAGGAACAGTGCAGCCTGCCTCACAAGTGCCATTCCAATCCACCCTCACAGCA
ACCTGGTGGAATTGTTATTTATGACCTTTTCTTACAAATGAGATTTCTGAAGCTCAGAGAAATTAAGCAACGAG
ATGAAGGTACCCAGCTGTGTGCACTGACCTGTTTGTAGAAAATACTGGCCTTCTGGGACCAAGGCAGGGATGCTT
TGCCCTGCCCTCTATGCTCTCTGTGCTCTCCACTCCCTCTCCCTCCTCCAACATTCCCTCCCTCTGTCTCC
AGCAGCCCCAGAGACCAGAATGATCCAGAGCTGGAGAAAGAACCCGAAGGCTCTTAGGGAGCAGCCAGAGGGCC
AAGTGACCAAGAGGATGGGGCTGAGCTGGGGAAGGGTGGCATCGAGGACCTTCTTGCACTTCTCTGTGGGAAG
CCCAGTGCCCTTGTCTCTGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT
GAGAGGCAGCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCTGCTGCCCTGCCACCCTCGGCCTCTGTCTGGAA
GCCCCACCCCTTTCTTCTGTACATAATGTCACTGGCTGTGTTGGGATTTTAAATTTATCTTCACTCAGCACCAAG
GGCCCCGACACTCCACTCCTGCTGCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTTGTATTTATTAATAAC
ATTTCTTTTTTCACTCTTTGGGCATGAGGTTGGCTCTTTGTGGCCAGGAACCTGAGTGGGGCCTGGTGGAGAAGGG
GCNAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
CGTGGCNCNTTGGCTGGCATNCCTGGGTTCCGCGAGAGGGGCTGGGGATGGTCTTGTGAGATGGTCTAGAGACTCAAG
AATTTAGGAAGTAGAAGCAGGATTTTGACTCAAGTTTGTAGTTTCCACATCGCTGGCCTGTTTGTGACTTCATG
TTTGAAGTTGCTCCAGAGAGAAATCAAAGGTGTCAACAGCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTTCT
TTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC

227/237

FIGURE 227

GGCCGAGCGGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTTCGGA
GGAACCCCTGCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAACTGCCCCCT
CCTTCTTTCTTCTTTTCTTCCCTCCCTCCCTTTCTTCCCCTTTTCTTCCCTTCCCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCCCTGAAGAGATTGCATTTGAGCCAGGTCCTGTAG

228/237

FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCCGAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCCGCTTCCGCGTGCTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTCACCCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGGCTACCCCCACCTCACCCCTGTGCTCCTTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGATATGTGTAGCGGGTACTGGTTCCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

229/237

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCCAAAACACCGAAAAACAAAAGAGAGAAACAACACCCCAACAACCTGGGGTGG
GGGGAAGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGCTCGAACCAGTGGCTGGCGGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCTCCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCCTGGGCGGCCGTGG
ACAACATGATGGTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTTTGCGGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGGCCACTGGGAAACCAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCCATCAGCAAAACCATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGGAATATGAATGCAGTGCGGAAAATGCTGTGTCAAT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACTTTGCTCCTACTATTTCAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAATTTTAGCACAAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCCCT
CCAAGTACAGCCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTCTCCTGCTGGTACCT
TGTGTTGACACTGTCCTCTTTCACCAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAAATACTTTAATTCTACTCTTTTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAAATCTACAACTTTTCAAAGCATTTCAGTCATGGT
CTGCTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATTTCAGTTCCTGTTCATAAGAGTCAATCCACATTTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCAATGATATTTCAGGTTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGTAAGTTTCAGCATGGGTGTGCCTTCATAC
AATAATATTTTTCTCTTTGTCTCCAATAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCCTGCACTGACAATGGGGTTTGAGAATTTTGGCCACACT
AACTCAGTTCTTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTTAAAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTTTCCTTTCACTAGTCCAAGCCAAAA
TTTTAAGATGATTTGTCAGAAAGGGCACAAAGTCCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAAATTTTATTTTGTGGCAGGTATTATGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACCTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAAGGTTTTTGG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCCATCTTTTTTTTTTAAATTTTTTCTGCCTATTTAA
AGACAAATTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGD TAVLR CYLED
GASKGAWLNRSSIIIFAGGDKWSVDPRVSISTLNKRDYSLQIQNV DVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVKV VVNFAPT IQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILT VTNVTQEHFGNYTCVAANKLGTTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLT LSSFTSIFYLKNAILO

Important features of the protein:**Signal peptide:**

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325**Myelin P0 protein:**

amino acids 92-121

237/237

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAACAAGCTGCTTTTACCAAATTGCAATGGAGCCTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTTCACTCAGAACCCAGGGATGTTTCAAGATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCAAGG
CCCACCGGGAGTCAAGGGGAGAGGCGGGCCTCCAAGGACCCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAAGGGGAACTGGAACCTAAGGGAGAGAAAGGAGACCTGGGTCTCCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCCTCCTGGAGCCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCCACCAGGTTTGGCTGGTFTTCTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCCGGGGCCCTCCTGGTGCAGTGGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCGAGATCTGGCTGGATAAT
GTTCAGTGTCGGGGCACGGAGAGTACCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGAATGTCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGGAAACCCTTTCA
CTTCTCTGCTCCCGAGGTGTCCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPSGPQGPVGKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGKGE
KGDGLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPGAKGDQGPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGTGLQGQQGRKGESGVP
GPA
GVKGEQGSPLAGPKGAPGQAGQKGDQGVKSSGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTTCATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCCAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTTAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTACTG
CACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTTCTTTGGACCAAAGATCTCTTTTCG
TGATTCCTTGCAACAATCAATTGAGAATCTTCATGTATTCTGGAGAACACCATTCCCTGATTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

234/237

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGATTGAGAGAGACTTTACCCCGCCGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
TGGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGTGGCGGGTGGCTTCTTTCTCCTCGGCTTCCTCTTCGGGTGGTTTA
TAAAATCCTCCAATGAAGCTACTAATTAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
GAATTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAACAAAACCTTTAGCTTGCAAAGCAAATCAATCCAGTGGAAGAATTTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGTCCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTAGTGCTTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACT
CCGACCCTGCTGACTACTTTGCTCCTGGGGTGAAGTCCTATCCAGACGGTTGGAATCTTCCT
GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAAATTGCAGAGGGCTGTTGGTCTTCCAA
GTATTCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCACCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
CTTTACTGGAACTTTTCTACACAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
CTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTCTACTGAGTGGGCA
GAGGAGAATTCAAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAACTACACTCTGAGAGTTGATTGTACACCGCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAAGTTGG
ACTAAAAAAGTCTTCCCCAGAGTTCAAGTGGCATGCCCAGGATAAGCAAATTGGGATCTGG
AAATGATTTTGAGGTGTTCTTCCAACGACTTGGAAATTGCTTCAGGCAGAGCACGGTATACTA
AAAATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTG
AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTGATTGTGAGATTATG
CTGTAGTTTAAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTCAAGTGAAGAGACTCCAGGACTTTGACAAAAGCAACCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTATTTGATCCATTAGGGTTACCAGAC
AGGCCTTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTTATGATGCTCTGTTTGATATTGAAAGCAAAGTGGACCCTTCCAAGGCC
GGGGAGAAAGTGAAGAGACAGATTTATGTTGCAGCCTTACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAAATTGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWNLLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEPPIPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVARIYRGKVFRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANAYARRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPFSGMPRISKLGSGNDFEVFFQRLGASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSFAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713

237/237

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
14 September 2000 (14.09.2000)

PCT

(10) International Publication Number
WO 00/53756 A3

(51) International Patent Classification⁷: **C12N 15/12**,
15/63, 5/10, 1/21, 1/19, C07K 14/47, 14/705, 19/00, 16/18,
16/28, G01N 33/68, 33/543, A61K 47/48, 51/08, 38/17

(21) International Application Number: **PCT/US00/04341**

(22) International Filing Date: 18 February 2000 (18.02.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PCT/US99/05028	8 March 1999 (08.03.1999)	US
60/123,957	12 March 1999 (12.03.1999)	US
60/126,773	29 March 1999 (29.03.1999)	US
60/130,232	21 April 1999 (21.04.1999)	US
60/131,445	28 April 1999 (28.04.1999)	US
60/134,287	14 May 1999 (14.05.1999)	US
60/141,037	23 June 1999 (23.06.1999)	US
60/145,698	26 July 1999 (26.07.1999)	US
60/162,506	29 October 1999 (29.10.1999)	US
PCT/US99/28313	30 November 1999 (30.11.1999)	US
PCT/US99/28551	2 December 1999 (02.12.1999)	US
PCT/US99/28565	2 December 1999 (02.12.1999)	US
PCT/US99/30095	16 December 1999 (16.12.1999)	US
PCT/US99/31243	30 December 1999 (30.12.1999)	US
PCT/US99/31274	30 December 1999 (30.12.1999)	US
PCT/US00/00219	5 January 2000 (05.01.2000)	US
PCT/US00/00277	6 January 2000 (06.01.2000)	US
PCT/US00/00376	6 January 2000 (06.01.2000)	US

(71) Applicant (for all designated States except US): **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ASHKENAZI, Avi, J.** [US/US]; 1456 Tarrytown Street, San Mateo, CA 94402 (US). **BAKER, Kevin, P.** [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). **BOTSTEIN, David**

[US/US]; 2539 Somerset Drive, Belmont, CA 94002 (US). **DESNOYERS, Luc** [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). **EATON, Dan, L.** [US/US]; 75 Knight Drive, San Rafael, CA 94901 (US). **FERRARA, Napoleone** [US/US]; 2090 Pacific Avenue #704, San Francisco, CA 94109 (US). **FILVAROFF, Ellen** [US/US]; 538 18th Avenue, San Francisco, CA 94121 (US). **FONG, Sherman** [US/US]; 19 Basinside Way, Alameda, CA 94502 (US). **GAO, Wei-Qiang** [CN/US]; 641 Pilgrim Drive, Foster City, CA 94404 (US). **GERBER, Hanspeter** [CH/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **GERRITSEN, Mary, E.** [CA/US]; 541 Parrott Drive, San Mateo, CA 94402 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GODOWSKI, Paul, J.** [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). **GRIMALDI, Christopher, J.** [US/US]; 1434 36th Avenue, San Francisco, CA 94122 (US). **GURNEY, Austin, L.** [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). **HILLAN, Kenneth, J.** [GB/US]; 64 Seward Street, San Francisco, CA 94114 (US). **KLJAVIN, Ivar, J.** [US/US]; 3673 Crescent Drive, Lafayette, CA 94549 (US). **KUO, Sophia, S.** [US/US]; Apartment 3, 59 Surrey Street, San Francisco, CA 94131 (US). **NAPIER, Mary, A.** [US/US]; 1015 Hayne Road, Hillsborough, CA 94010 (US). **PAN, James** [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). **PAONI, Nicholas, F.** [US/US]; 1756 Terrace Drive, Belmont, CA 94002 (US). **ROY, Margaret, Ann** [US/US]; 2960 Webster Street #4, San Francisco, CA 94123 (US). **SHELTON, David, L.** [US/US]; 5845 Clover Drive, Oakland, CA 94618 (US). **STEWART, Timothy, A.** [US/US]; 465 Douglass Street, San Francisco, CA 94114 (US). **TUMAS, Daniel** [US/US]; 3 Rae Avenue, Orinda, CA 94563 (US). **WILLIAMS, P., Mickey** [US/US]; 509 Alto Avenue, Half Moon Bay, CA 94019 (US). **WOOD, William, I.** [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US).

(74) Agents: **KRESNAK, Mark, T. et al.**; Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080-4990 (US).

(81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

[Continued on next page]

(54) Title: **SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME**

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 00/53756 A3



(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(88) Date of publication of the international search report:
1 February 2001

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *With international search report.*

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/04341

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7	C12N15/12	C12N15/63	C12N5/10	C12N1/21	C12N1/19
	C07K14/47	C07K14/705	C07K19/00	C07K16/18	C07K16/28
	G01N33/68	G01N33/543	A61K47/48	A61K51/08	A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 57983 A (ZYMOGENETICS INC) 23 December 1998 (1998-12-23) the whole document	1-3, 5-13, 15-21
A	KLEIN R D ET AL: "Selection for genes encoding secreted proteins and receptors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, no. 93, 1 July 1996 (1996-07-01), pages 7108-7113, XP002077277 ISSN: 0027-8424 the whole document	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

20 July 2000

Date of mailing of the international search report

18. 10. 00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/04341

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YOKOYAMA-KOBAYASHI M ET AL: "A signal sequence detection system using secreted protease activity as an indicator" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 163, no. 2, 3 October 1995 (1995-10-03), pages 193-196, XP004041983 ISSN: 0378-1119 the whole document ----	
A	EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document ----	
A	WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document ----	
P,X, L	WO 99 46281 A (BAKER KEVIN P ;CHEN JIAN (US); GENENTECH INC (US); GURNEY AUSTIN () 16 September 1999 (1999-09-16) whole document, particularly the passages referring to PR0213 (e.g. pages 2,50,123,183), and the claims ----	1-3, 5-13, 15-21
P,X	WO 99 54437 A (MILLENNIUM BIOTHERAPEUTICS INC) 28 October 1999 (1999-10-28) the whole document -----	1-3, 5-13, 15-21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/04341

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Invention 1 : claims 1-3, 5-13, 15-21, all partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 1: claims 1-3,5-13,15-21, all partially

Nucleic acid with seq.ID.1, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.2 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.2 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide.

Invention 2: claims 1-29,38-43, and 50-53,
all partially

Nucleic acid with seq.ID.611, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.612 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.612 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PR0337 (seq.ID.523) using its interaction with PR04993 (seq.ID.612), method for linking a bioactive molecule to a cell expressing PR0337 through the use of PR04993, and method of modulating at least one activity of said cell thereby.

Invention 3: claims 1-29,38-43, and 50-53,
all partially

Nucleic acid with seq.ID.522, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.523 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.523 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PR04993 (seq.ID.612) using its interaction with PR0337 (seq.ID.523), method for linking a bioactive molecule to a cell expressing PR04993 through the use of PR0337, and method of modulating at least one activity of said cell thereby.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Invention 4: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.613, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.614 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.614 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PR0700 (seq.ID.523), PR0725 (seq.ID.616) and/or PR0739 (seq.ID.618) using their interaction with PR01559 (seq.ID.614), method for linking a bioactive molecule to a cell expressing PR0700, PR0725 or PR0739 through the use of PR01559, and method of modulating at least one activity of said cell thereby.

Invention 5: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.89, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.90 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.90 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of said peptide fused to a heterologous sequence, isolated extracellular domain of said protein or said protein lacking its signal peptide, and an antibody against said polypeptide. Also a method of detecting PR01559 (seq.ID.614) using its interaction with PR0700 (seq.ID.90), method for linking a bioactive molecule to a cell expressing PR01559 through the use of PR0700, and method of modulating at least one activity of said cell thereby.

Invention 6: claims 1-21,30-37,44-49, and 54-57,
all partially

Nucleic acid with seq.ID.615, encoding a polypeptide comprising the amino acid sequence as represented in seq.ID.616 or a nucleic acid having at least 80% homology thereto, vector comprising said nucleic acid, host cell comprising said vector, process for producing the protein of seq.ID.616 using said host, the isolated protein or one having at least 80% homology thereto, a chimeric protein of

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO381, as represented by seq.ID's 144 and 145,
PRO386, as represented by seq.ID's 149 and 150,
PRO540, as represented by seq.ID's 156 and 157,
PRO615, as represented by seq.ID's 161 and 162,
PRO618, as represented by seq.ID's 168 and 169,
PRO719, as represented by seq.ID's 177 and 178,
PRO724, as represented by seq.ID's 182 and 183,
PRO772, as represented by seq.ID's 189 and 190,
PRO852, as represented by seq.ID's 195 and 196,
PRO853, as represented by seq.ID's 205 and 206,
PRO860, as represented by seq.ID's 210 and 211,
PRO846, as represented by seq.ID's 215 and 216,
PRO862, as represented by seq.ID's 220 and 221,
PRO864, as represented by seq.ID's 225 and 226,
PRO792, as represented by seq.ID's 230 and 231,
PRO866, as represented by seq.ID's 235 and 236,
PRO871, as represented by seq.ID's 244 and 245,
PRO873, as represented by seq.ID's 253 and 254,
PRO940, as represented by seq.ID's 258 and 259,
PRO941, as represented by seq.ID's 263 and 264,
PRO944, as represented by seq.ID's 269 and 270,
PRO983, as represented by seq.ID's 283 and 284,

Claim : 9

PRO1057, as represented by seq.ID's 295 and 296,
PRO1071, as represented by seq.ID's 300 and 301,
PRO1072, as represented by seq.ID's 302 and 303,
PRO1075, as represented by seq.ID's 308 and 309,
PRO181, as represented by seq.ID's 321 and 322,
PRO195, as represented by seq.ID's 329 and 330,
PRO865, as represented by seq.ID's 336 and 337,
PRO827, as represented by seq.ID's 345 and 346,
PRO1114, as represented by seq.ID's 351 and 352,
PRO237, as represented by seq.ID's 357 and 358,
PRO541, as represented by seq.ID's 362 and 363,
PRO273, as represented by seq.ID's 369 and 370,
PRO701, as represented by seq.ID's 374 and 375,
PRO704, as represented by seq.ID's 379 and 380,
PRO706, as represented by seq.ID's 384 and 385,
PRO707, as represented by seq.ID's 389 and 390,
PRO322, as represented by seq.ID's 394 and 395,
PRO526, as represented by seq.ID's 399 and 400,
PRO531, as represented by seq.ID's 404 and 405,
PRO534, as represented by seq.ID's 409 and 410,
PRO697, as represented by seq.ID's 414 and 415,
PRO717, as represented by seq.ID's 419 and 420,
PRO731, as represented by seq.ID's 424 and 425,
PRO218, as represented by seq.ID's 429 and 430,
PRO768, as represented by seq.ID's 436 and 437,
PRO771, as represented by seq.ID's 441 and 442,
PRO733, as represented by seq.ID's 446 and 447,
PRO162, as represented by seq.ID's 451 and 452,
PRO788, as represented by seq.ID's 453 and 454,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

PRO1008, as represented by seq.ID's 455 and 456,
PRO1012, as represented by seq.ID's 458 and 459,
PRO1014, as represented by seq.ID's 463 and 464,
PRO1017, as represented by seq.ID's 465 and 466,
PRO474, as represented by seq.ID's 467 and 468,
PRO1031, as represented by seq.ID's 469 and 470,
PRO938, as represented by seq.ID's 471 and 472,
PRO1082, as represented by seq.ID's 476 and 477,
PRO1083, as represented by seq.ID's 482 and 483,
PRO200, as represented by seq.ID's 487 and 488,
PRO285, as represented by seq.ID's 495 and 496,
PRO286, as represented by seq.ID's 497 and 498,
PRO213-1, as represented by seq.ID's 505 and 506,
PRO1330, as represented by seq.ID's 507 and 508,
PRO1449, as represented by seq.ID's 509 and 510,
PRO298, as represented by seq.ID's 514 and 515,
PRO403, as represented by seq.ID's 525 and 526,

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/04341

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9857983	A	23-12-1998	AU 7979898 A EP 0996628 A	04-01-1999 03-05-2000
EP 0834563	A	08-04-1998	JP 10179178 A US 5824504 A	07-07-1998 20-10-1998
WO 9707198	A	27-02-1997	US 5707829 A AU 6712396 A AU 6768596 A CA 2227220 A CA 2229208 A EP 0839196 A EP 0851875 A JP 11510045 T US 6043344 A WO 9704097 A US 6074849 A US 5969093 A	13-01-1998 18-02-1997 12-03-1997 06-02-1997 27-02-1997 06-05-1998 08-07-1998 07-09-1999 28-03-2000 06-02-1997 13-06-2000 19-10-1999
WO 9946281	A	16-09-1999	AU 3072199 A AU 3075099 A WO 9947677 A AU 1532499 A EP 1032667 A WO 9927098 A AU 3757099 A WO 9954467 A AU 1070399 A EP 1025227 A WO 9920756 A	27-09-1999 11-10-1999 23-09-1999 15-06-1999 06-09-2000 03-06-1999 08-11-1999 28-10-1999 10-05-1999 09-08-2000 29-04-1999
WO 9954437	A	28-10-1999	AU 3965799 A	08-11-1999